UNIVERSIDAD CARLOS III DE MADRID

MATHEMATICS DEPARTAMENT

Ph.D. THESIS

CLASSICAL PERTURBATIONS FOR MATRICES OF LINEAR FUNCTIONALS

AUTHOR:

Juan Carlos García Ardila

DIRECTOR:

Dr. Francisco Marcellán Español

Leganés, Spain, July 2017

UNIVERSIDAD CARLOS III DE MADRID

Ph.D. THESIS

CLASSICAL PERTURBATIONS FOR MATRICES OF LINEAR FUNCTIONALS

AUTHOR: Juan Carlos García Ardila

DIRECTOR: Dr. Francisco Marcellán Español

Firma del tribunal calificador:

		Firma
Presidente:		
Vocal:		
Secretario		

Calificación

Leganés, de de 201

AGRADECIMIENTOS.

Quiero agradecer al profesor Francisco Marcellán por guiarme y brindarme su consejo siempre que lo necesite durante los casi 5 años que he estado en la UC3M. A las dos mujeres que más amo en la vida, mi madre y la Crespa, las cuales siempre han estado a mi lado brindándome su apoyo y soporte, a ellas va dedicada esta memoria. A mis compañeros de despacho (los pichus), que más que eso se han convertido en parte de mi familia, Misa, Ernesto y David les agradezco de todo corazón todos los buenos momentos vividos y su apoyo en los momentos difćiles. Quiero también agradecer a Maria, Pablo y a los dos Ignacios por las comidas tan agradables y enriquecedoras en estos años y a Fredy (el asere) por las inolvidables cenas junto a su maravillosa familia. Por último y no menos importantes quiero agradecer a Herbert, Edy, Amilcar, Luis, Natalia, y tantos otros que en este momento se me olvidan pero que han sido parte de este proceso, a todos ustedes !muchas gracias!

Resumen

El objetivo de esta Tesis es estudiar trasformaciones espectrales para matrices que tiene como entradas funcionales lineales. En particular estudiamos las transformaciones de Christoffel, Geronimus y Geronimus-Uvarov. Con el fin de que esta Tesis sea lo más autocontenida posible, la hemos dividido en siete capítulos.

- En el **Capítulo 1** introducimos algunos conceptos y fijamos la notación que será usada a lo largo de esta Tesis a la vez que exponemos algunas propiedades básicas acerca de matrices semi-infinitas, módulos y polinomios matriciales junto con su teoría espectral. Hecho esto, pasamos a introducir la definición de matriz de funcionales lineales (y su forma sesquilineal asociada) también como el concepto de bi-ortogonalidad.

- En el **Capítulo 2** resumimos algunos resultados relativos a formas bilineales para las cuales el operador multiplicación por un polinomio es simétrico y su conexión con las relaciones de recurrencia de orden superior [51]. Con esto en mente, pasamos a explicar cuidadosamente el resultado principal de [64] concerniente a la relación existente entre una sucesión de polinomios satisfaciendo una relación de recurrencia de orden superior y una sucesión de polinomios matriciales que son ortogonales respecto a una medida matricial. Para concluir, mostramos algunos ejemplos de gran interés en la literatura.

- El **Capítulo 3** está dividido en dos partes. La primera parte se basada en los trabajos de A. J. Durán [51], A. J. Durán y W. Van Assche [64], y M. Derevyagin and F. Marcellán [48] siendo este último el referente para esta primera parte del capítulo. Aquí, damos algunos resultados relacionados con la transformación espectral de Geronimus en el contexto de funcionales lineales y con esto en mente, motivamos la definición de este tipo de transformaciones para las formas bilineales simétricas $B(\cdot,\cdot)$ definidas en términos de una medida de probabilidad con la propiedad de que el operador multiplicación por un cierto polinomio $h(x)$ es simétrico para la forma bilineal, esto es $B(hf,g) = B(f,gh)$ para todo $f,g \in \mathbb{R}[x]$. Acto seguido, introducimos la noción de transformación de Geronimus múltiple y generalizamos los resultados dados en [109], encontrando, por ejemplo, que los productos internos tipo Sobolev discretos obtenidos al aplicarle un transformación de Geronimus múltiple a una forma bilineal nos lleva a una transformación de Geronimus matricial. Estos resultados han sido publicados en [47].

  Motivados por el resultado anterior, en la segunda parte de este capítulo, estudiamos la transformación de Geronimus (hermitica) pero ahora, sobre una matriz de medidas definida positiva, es decir, estamos interesados en el análisis de formas sesquilineales $\langle \cdot,\cdot \rangle_W$ defindas mediante

  $$\langle P(x)W(x), Q(x)W(x)\rangle_W = \int P(x)dMQ^{\dagger}(x),$$

  donde M es una matriz de medidas definida positiva y $W(x)$ es un polinomio matricial de grado fijo pero arbitrario. Aquí, encontramos condiciones para la existencia de la sucesión

de polinomios ortogonales respecto a $\langle\cdot,\cdot\rangle_W$ así como la fórmula de conexión entre los polinomios ortogonales originales y perturbados. Los resultados de este capítulo han sido publicados en [67, 68].

- El **Capítulo 4** se dedica a la extensión de la fórmula de Christoffel para polinomios bi-ortogonales matriciales. Más precisamente, dada una matriz de funcionales lineales u y $W(x)$ un polinomio matricial con coeficiente principal no-singular, nosotros estudiamos la siguiente transformación matricial $\hat{u} = W(x)u$ y la relación entre sus familias de polinomios bi-ortogonales. El resultado principal de este capítulo es precisamente el Teorema 4.11, que presenta las fórmulas de conexión entre las sucesiones de polinomios bi-ortogonales matriciales originales y perturbados. Para este propósito usamos toda la riqueza de la teoría espectral disponible para polinomios matriciales, en particular las cadenas de Jordan y los "roots polynomials" a izquierda y derecha que serán extremadamente útiles [77, 110]. Finalmente, veremos que estas transformaciones de Christoffel se pueden extender a la teoría de sistemas integrables. Los resultados de este capítulo han sido publicados en [11].

- El **Capítulo 5** estudia la extensión de la transformación de Geronimus para matrices de funcionales lineales soportados en la recta real, es decir, multiplicaremos una matriz de funcionales lineales por un polinomio matricial $W_G(x)$ y le adicionaremos una suma de masas adecuadas (que dependen de los "roots polynomials" a izquierda y derecha). Aquí, desarrollamos dos diferentes métodos, el espectral y no espectral con el fin de obtener fórmulas de conexión entre las sucesiones de polinomios bi-ortogonales matriciales asociados al funcional original y el perturbado.

- En el **Capítulo 6** se desarrolla la extensión de las transformación de Geronimus-Uvarov para matrices de funcionales lineales soportados sobre la recta real. Este tipo de transformaciones pueden considerarse como una composición de una transformación de Geronimus y, acto seguido, de una transformación de Christoffel. En términos de matrices de funcionales esto se escribe como $u \to \check{u} \to \hat{u}$, donde $\check{u}W_G(x) = u$ y $\hat{u} = W_C(x)\check{u}$, con $W_C(x)$ y $W_G(x)$ polinomios matriciales. Como en el capítulo 5, usando el método espectral y no-espectral, obtenemos fórmulas de conexión entre los polinomios bi-ortogonales matriciales originales y los perturbados. En el método espectral encontramos una representación para los polinomios bi-ortogonales matriciales perturbados en términos de la sucesiones de polinomios bi-ortogonales originales y las funciones de segunda especie (ver (1.9)). Aquí asumimos que los coeficientes principals de los polinomios $W_C(x)$ y $W_G(x)$ son matrices no-singulares. En el método no-espectral damos una representación para los polinomios bi-ortogonales matriciales perturbados sin asumir ninguna hipótesis sobre el coeficiente principal del polinomio $W_G(x)$. Finalmente, como una aplicación, estudiamos la transformación de Uvarov matricial, que consiste en adicionarle al funcional original una suma de masas. Los resultados del capítulo 5 y 6 han sido publicados en [12].

- Finalmente, en el **Capítulo 7** damos un resumen de los principales resultados de esta tesis, así como una lista de problemas abiertos.

Table of Contents

Introduction

Historical notes

The theory of linear functionals and their connection with the theory of orthogonal polynomials is well known in the literature. Several examples of perturbations of a linear functional $\mathbf{u}$ have been studied (see [10, 7, 22, 26, 27, 34, 76, 115, 116, 148]). In particular, when we deal with the positive definite case, i.e. the linear functional has an integral representation in terms of a nontrivial probability measure supported in a infinity subset of the real line, such perturbations provide an interesting information in the framework of Gaussian quadrature formulas taking into account the perturbation yields new nodes and Christoffel numbers. Three canonical perturbations have attracted the interest of researchers.

The Christoffel transformation that appears when a linear functional $\mathbf{u}$ is transformed into $\hat{\mathbf{u}}_1 = w_c(x)\mathbf{u}$ with $w_c(x) \in \mathbb{R}[x]$ was studied for first time in [36]. Here, Christoffel studies a quadrature problem for the Lebesgue measure supported in $[-1,1]$, but now preassigning a certain number of nodes outside the open set $(-1,1)$. The above theory is easily extended to weighted integrals over a interval $[a,b]$ (see [69]). Thus, if $I(f) = \int_a^b f(x)d\mu$, with $\mu'(x) \geq 0$, he deals with an approximation of this integral by means of the following quadrature formula

$$I(f) = Q_n(f) + R_n(f), \quad \text{with} \quad Q_n(f) = \sum_{k=1}^{n} \lambda_{n,k} f(x_{n,k}) + \sum_{m=1}^{N} u_{n,m} f(\gamma_{n,m}),$$

being $R_n(f)$ the remainder, $\gamma_{n,1}, \ldots, \gamma_{n,N}$, fixed numbers outside the interval (a,b) and $x_{n,1}, \ldots, x_{n,n}$, free nodes. The weights $\lambda_{n,k}$ and $u_{n,m}$ are assumed to be free variables. If now we define

$$\hat{p}_n(x) = \prod_{k=1}^{n}(x - x_{n,k}) \quad \text{and} \quad w_c(x) = \prod_{m=1}^{N} |x - \gamma_{n,m}|,$$

then the quadrature formula has maximal degree of exactness $(2n-1+N)$ if and only if

$$I(\hat{p}_n(x) w_c(x) q(x)) = 0 \quad \text{for every} \quad q(x) \in \mathbb{R}_{n-1}[x].$$

The above implies that the quadrature formula has maximal degree of exactness if $\hat{p}_n(x)$ is orthogonal to every polynomial of degree $\leq n-1$ with respect to to the measure $d\hat{\mu} = w_c(x)d\mu$ and

$$\lambda_{n,k} = I\left[\frac{w_c(\cdot)\hat{p}_n(\cdot)}{w_c(x_{n,k})\hat{p}_n'(x_{n,k})(\cdot - x_{n,k})}\right], \quad u_{n,k} = I\left[\frac{w_c(\cdot)\hat{p}_n(\cdot)}{w_c'(\gamma_{n,m})\hat{p}_n(\gamma_{n,m})(\cdot - \gamma_{n,m})}\right],$$

with $m = 1, \cdots, N$, and $k = 1, \cdots, n$. Moreover, Christoffel obtains an explicit representation of the polynomial $\hat{p}_n(x)$ in terms of the sequence of orthogonal polynomials $(p_n(x))_{n\in\mathbb{N}}$ with respect to the measure $d\mu$ as follows (see [69])

$$w_c(x)p_n(x) = \text{const} \cdot \det \begin{bmatrix} p_n(x) & p_{n+1}(x) & \cdots & p_{n+N}(x) \\ p_n(\gamma_{n,1}) & p_{n+1}(\gamma_{n,1}) & \cdots & p_{n+N}(\gamma_{n,1}) \\ \vdots & \vdots & \cdots & \vdots \\ p_n(\gamma_{n,N}) & p_{n+1}(\gamma_{n,N}) & \cdots & p_{n+N}(\gamma_{n,N}) \end{bmatrix}.$$

This is now commonly refereed as Christoffel theorem. More recently, the Christoffel transformation has been studied for linear functionals (independently of the existence of an integral representation). For example in [33, 148] necessary and sufficient conditions for the functional $\hat{\mathbf{u}}$ be quasi-definite are given, while in [138, 149] a connection formula between the sequences of orthogonal polynomials with respect to $\mathbf{u}$ and $\hat{\mathbf{u}}$ is deduced.

The Geronimus transformation appears when you are dealing with perturbed functionals $\check{\mathbf{u}}$ defined by $w_g(x)\check{\mathbf{u}} = \mathbf{u}$, where $w_g(x)$ is a polynomial. This class of transformations was first discussed by J. Shohat [135] for a particular case of linear functionals. Given a sequence of orthogonal polynomials $p_n(x)$ with respect to a measure supported in the real line, find necessary and sufficient conditions on the sequence of real numbers $(a_n)_{n\in\mathbb{N}}$ in order to the sequence of polynomials $(Q_n(x))_{n\in\mathbb{N}}$ defined by $Q_n(x) := p_n(x) + a_n p_{n-1}(x)$, $n \geq 1$, is orthogonal with respect to some measure supported in the real line. However, he only gave a partial answer by using Favard's Theorem. Few years after Shohat's publication, the problem was studied in an exhaustive way by Geronimus [76] in the framework of the Hahn characterization of classical orthogonal polynomials (Hermite Laguerre, Jacobi, and Bessel). Indeed, to find necessary and sufficient conditions in order to the sequences $(p_n(x))_{n\in\mathbb{N}}$ and $(\frac{p_{n+1}'(x)}{n+1})_{n\in\mathbb{N}}$ be orthogonal. In the same way as for the Christoffel transformation, the sequence of orthogonal polynomials associated with the functional $\check{\mathbf{u}}$ has a determinantal representation but now in terms of the original polynomials and the second kind functions which are defined as a formal series, $C_n(z) =: \sum_{k=n}^{\infty} \frac{\langle \mathbf{u}, p_n(x)x^k \rangle}{z^{k+1}}$, see for example [115, 149].

The more general problem related to linear functionals $\mathbf{u}$ and $\tilde{\mathbf{u}}$ satisfying $w_g(x)\tilde{\mathbf{u}} = w_c(x)\mathbf{u}$, (where again $w_c(x)$ and $w_g(x)$ are polynomials) has been analyzed by Uvarov in [143] in the framework of measures. In the first part of paper, he studied the following problem. Given a measure μ supported in the real line and two sets of real numbers $\{\alpha_1, \cdots \alpha_l\}$ and $\{\beta_1, \cdots \beta_k\}$ with β_j, $j = 1, \ldots, k$, outside of the support of the measure, what is the relation between the sequence of orthogonal polynomials $(p_n(x))_{n\in\mathbb{N}}$ with respect to μ and the sequence of orthogonal polynomials $(p_n^{(k,l)}(x))_{n\in\mathbb{N}}$ with respect to the new measure defined by $d\tilde{\mu} = \prod_{j=1}^{l}(x-\alpha_j)\left[\prod_{j=1}^{k}(x-\beta_j)\right]^{-1} d\mu$?. Here, he finds the following determinantal formula

$$p_n^{(n,k)}(x) = const. \left[\prod_{j=1}^{k}(x-\beta_j)\right]^{-1} \det \begin{bmatrix} p_{n-k}(\alpha_1) & \cdots & p_{n+l}(\alpha_1) \\ \cdots & \cdots & \cdots \\ p_{n-k}(\alpha_l) & \cdots & p_{n+l}(\alpha_l) \\ C_{n-k}(\beta_l) & \cdots & C_{n+l}(\beta_l) \\ \cdots & \cdots & \cdots \\ C_{n-k}(\beta_k) & \cdots & C_{n+l}(\beta_k) \\ p_{n-k}(x) & \cdots & p_{n+l}(x) \end{bmatrix}.$$

In the second part, he deals with the addition of a finite number of Dirac masses i.e., he studies the sequence of orthogonal polynomials with respect to the measure $d\tilde{\mu} = d\mu + \sum_{j=1}^{s} m_j \delta(x-x_j)$ where $\delta(x)$ is the Dirac delta supported in zero, m_j are positive constants, and x_j are real numbers (Uvarov transformation). This kind of transformations appear in the analysis of polynomial eigenfunctions of fourth order linear differential operators with polynomial coefficients (see [96]). In [34] necessary and sufficient conditions for the functional $\tilde{\mathbf{u}}$ be quasi-definite are given. In [138, 149] an expression of orthogonal polynomials with respect to $\tilde{\mathbf{u}}$ in terms of orthogonal polynomials with respect to $\mathbf{u}$ is obtained.

Notice that, in particular, if we take $w_c(x) = w_g(x)$, then a Christoffel transformation applied to Geronimus transformation, transforms $\check{\mathbf{u}}$ into the original linear functional $\mathbf{u}$ i.e. the Christoffel transformation is the left inverse of the Geronimus transformation. However, a Geronimus transformation applied to Christoffel transformation, transforms $\hat{\mathbf{u}}$ into $\tilde{\mathbf{u}}$, i.e. an Uvarov transformation. The above three transformations are known in the literature as Darboux transformations. They have appeared in the framework of the bispectral problem (see [80]).

In the last years, spectral properties of the monic Jacobi matrix associated with the sequence of monic orthogonal polynomials with respect to a linear functional $\mathbf{u}$ have been studied, finding a close relation between the perturbation of linear functionals (in particular for the above three transformations) and their LU and UL factorizations, where L is a lower triangular matrix and U is an upper triangular matrix (see [27, 148]). For example in [148] it is showed that if the monic Jacobi matrix J_{mon} associated with a quasi-definite linear functional $\mathbf{u}$ can be factorized as $J_{mon} = LU + aI$, $a \in \mathbb{R}$, then the monic Jacobi matrix with respect to the functional $\hat{\mathbf{u}}$ defined as above (with $w_c = (x-a)$) is given by $UL + aI$ (Darboux transformation without parameter of J_{mon}). In the same way, if J_{mon} can be factorized as $J_{mon} = UL + aI$, $a \in \mathbb{R}$, then the monic Jacobi matrix with respect to the functional $\check{\mathbf{u}}$ defined previously (with $w_g = (x-a)$) is given by $LU + aI$ (Darboux transformation with parameter of J_{mon}). See also [27, 47].

A more general concept than a linear functional is the symmetric bilinear form. It is well known that every linear functional $\mathbf{u}$ yields a bilinear form B defined by $B(f,g) = \langle \mathbf{u}, fg \rangle$. Notice that the inverse is not true at all. One of the most well known positive definite symmetric bilinear forms are associated with the so called discrete Sobolev type inner products $\langle \cdot, \cdot \rangle_S$ which appear as a perturbation of a positive definite linear functional $\mathbf{u}$ as follows

$$\langle f,g\rangle_S = \langle \mathbf{u}, fg\rangle + \sum_{k=0}^{j} M_k f^{(k)}(\alpha) g^{(k)}(\alpha), \tag{1}$$

where $\alpha \in \mathbb{R}$, $M_k \geq 0, 0 \leq k \leq j$, and f,g are polynomials. The first paper dealing with Sobolev type inner products is due to D. C. Lewis in [103]. There, a problem of approximation of a function and its derivatives by algebraic polynomials using least-squares was studied. In [90, 91] R. Koekoek, and R. Koekoek and H. G. Meijer studied particular cases of (1), when the linear functional $\mathbf{u}$ is the Laguerre linear functional, while that in [8, 9, 117] some properties of the sequence of orthogonal polynomials with respect to the inner product (1) are studied. In a more general framework the following Sobolev type inner products have been studied (see [120])

$$\langle f,g\rangle_S = \sum_{k=0}^{j} M_k \left\langle f^{(k)}, g^{(k)} \right\rangle_k$$

with f,g polynomials, and

$$\langle f,g\rangle_k = \int_{\Omega_k} f(x)g(x)d\mu_k,$$

where μ_k, $k=1,\dots,j$, are positive Borel measures such that $\operatorname{supp}(\mu_k) \subseteq \Omega_k \subseteq \mathbb{R}$. From here, a lot of contributions have been given (see for example [6, 120]). Notice that, in general, the above Sobolev type inner product can not be induced by a linear functional because otherwise it would satisfy $\langle xf,g\rangle_S = \langle f,xg\rangle_S$ for every polynomial f,g.

An important concept that follows as a particular case from the above definition is the concept of coherent pair introduced by A. Iserles, P. E. Koch, S. P. Nørsett, and J. M. Sanz-Serna, (see [86]). This concept emerges to the study the following Sobolev type inner product

$$\langle f,g\rangle_S = \int_a^b fg\,d\mu_0 + \lambda \int_a^b f'g'\,d\mu_1, \tag{2}$$

where μ_0,μ_1 are positive Borel measures and $\lambda > 0$. If $(P_n(x))_{n\in\mathbb{N}}$ is the sequence of monic orthogonal polynomials with respect to $d\mu_0$ and $(T_n(x))_{n\in\mathbb{N}}$ is the sequence of the monic orthogonal polynomials with respect to $d\mu_1$, then the pair $(d\mu_0,d\mu_1)$ is said to be a coherent pair, if there exist nonzero constants $(\beta_n)_{n\in\mathbb{N}}$ such that

$$T_n(x) = \frac{P'_{n+1}(x)}{n+1} + \beta_n \frac{P'_n(x)}{n} \quad \text{for all} \quad n \geq 1.$$

If $(d\mu_0,d\mu_1)$ is a coherent pair, then the sequence of orthogonal polynomials with respect to (2) has an interesting structure (see [85, 118, 119]). Thus, it is relevant to investigate under what conditions $(d\mu_0,d\mu_1)$ constitute a coherent pair. As a consequence, a lot of contributions dealing with this problem have been given (see for example [85, 108, 118, 119]).

M. G. Krein in [97] (1949) was the first to discuss about of matrix orthogonal polynomials on the real line, when he was dealing with block Jacobi matrices, mainly as applied to self adjoint extension. Years later (1983) A. I. Aptekarev and E. M Nikishin [18] studied properties of Sturm-Liouville discrete operators which can be identified with a block Jacobi matrix

$$J = \begin{bmatrix} E_0 & D_1 & & & \\ D_1^\dagger & E_1 & D_2 & & \\ & D_2^\dagger & E_2 & D_3 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \tag{3}$$

where each matrix E_i and D_i has size $N \times N$, with E_i a Hermitian matrix and D_i a nonsingular matrix. They proved that there exists a positive definite matrix of measures of size $N \times N$ such that the sequence of matrix polynomials $(Q_n(x))_{n\in\mathbb{N}}$ defined recursively as

$$\begin{aligned} &D_{k+1}Q_{k+1}(x) = (xI_N - E_k)Q_k(x) - D_k^\dagger Q_{k-1}(x), \quad k > 1, \\ &Q_0(x) = I_N, \qquad E_1 Q_1 = (I_N - E_0), \end{aligned} \tag{4}$$

is a sequence of orthogonal polynomials with respect to such a matrix of measures.

The theory of matrix orthogonal polynomials was in a stand-by until that in 1993. A. J. Durán in [51] gave the characterization of symmetric bilinear forms for which the multiplication operator by a polynomial is symmetric. He deduces necessary and sufficient conditions for that a sequence of scalar polynomials $(p_n(x))_{n\in\mathbb{N}}$ satisfying a $(2N+1)$ recurrence relation

$$h(x)p_n(x) = c_{n,0}p_n(x) + \sum_{k=1}^{N} [\overline{c}_{n,k}p_{n-k}(x) + c_{n+k,k}p_{n+k}(x)],$$

is orthogonal with respect to a symmetric bilinear form (generalization of the Favard's theorem). In particular, he focused the attention on the discrete Sobolev type inner product. This work, gave a first idea about the relation between scalar orthogonal polynomials with respect to a bilinear form and matrix orthogonal polynomials with respect to a positive definite matrix of measures. Thus in (1995) A. J. Durán [52] showed that any sequence of scalar polynomials $(p_n(x))_{n\in\mathbb{N}}$ satisfying a $(2N+1)$ recurrence relation is orthogonal with respect to a positive definite matrix of measures of size $N \times N$. To prove the above statement is necessary identify each polynomial $p_n(x)$ with a certain vector $[R_{h,0}(p_n), \cdots, R_{h,N-1}(p_n)]$ (see Definition 2.9).

From the above result, A. J. Durán and W. Van Assche [64] proved that if $(p_n(x))_{n\in\mathbb{N}}$ is a sequence of scalar polynomials satisfying a $(2N+1)$ recurrence relation, then the sequence of matrix polynomials defined as

$$P_n(x) = \begin{bmatrix} R_{h,0}(p_{nN})(x) & R_{h,1}(p_{nN})(x) & \cdots & R_{h,N-1}(p_{nN})(x) \\ R_{h,0}(p_{nN+1})(x) & R_{h,1}(p_{nN+1})(x) & \cdots & R_{h,N-1}(p_{nN+1})(x) \\ \vdots & \vdots & & \vdots \\ R_{h,0}(p_{nN+N-1})(x) & R_{h,1}(p_{nN+N-1})(x) & \cdots & R_{h,N-1}(p_{nN+N-1})(x) \end{bmatrix}$$

satisfies a three term recurrence relation as in (4). The example (3.1) in [64] (see also the Example 2.11 in this work) exhibits the importance of the above result, since it shows that a discrete Sobolev type inner product can be always represented in terms of a matrix inner product when you add to the matrix of measures a mass point in the origin. The above approach opens the door to study spectral transformations for matrices of measures supported in the real line. This thesis gives an exhaustive study of the spectral transformations mentioned previously, based on an earlier work of M. Derevyagin and F. Marcellán (see [48]).

The above results reawaked the interest by matrix orthogonal polynomials. So, in [61] A. J. Durán and P. López-Rodríguez studied the properties of zeros (see Definition 1.11) of a sequence of matrix polynomials $(P_n(x))_{n\in\mathbb{N}}$ which are orthonormal with respect to a matrix of measures W that is positive definite, or equivalent to the sesquilinear form

$$\langle P,Q\rangle_W = \int P(x)dW(x)Q^{\dagger}(x) \quad P,Q \in \mathbb{C}^{N\times N}[x].$$

They showed that for every $n\in\mathbb{N}$, the zeros of the matrix polynomial $P_n(x)$ are precisely the eigenvalues of the block matrix J_n with the same multiplicity. Here J_n is the n-th block truncation of the block Jacobi matrix (3) associated to $(P_n(x))_{n\in\mathbb{N}}$. To difference of the scalar case, the zeros of $P_n(x)$ can not be simple, moreover, any two polynomials of the sequence $(P_n(x))_{n\in\mathbb{N}}$ can have a zero in common. The following theorem which is supported on the fact that for any matrix polynomial $A(x)$ with a zero a of multiplicity α

$$\frac{d^s}{dx}\operatorname{adj}(A(x))\Big|_{x=a} = 0_{N\times N}, \quad s=0,\cdots,\alpha-2, \quad \text{and,} \quad \frac{d^{\alpha-1}}{dx}\operatorname{adj}(A(x))\Big|_{x=a} \neq 0_{N\times N}$$

gives us some information about the zeros of $P_n(x)$.

Theorem 1. ([61]). i) *The zeros of $P_n(x)$ have a multiplicity not bigger than N. Furthermore $P_n(x)$ has nN zeros (counting multiplicities) and all zeros are real.*

ii) *If a is a zero of multiplicity α of $P_n(x)$, then $\operatorname{rank}(P_n(a)) = N-\alpha$. If a is a zero of $P_n(x)$ and $P_{n+1}(x)$, then $P_n(a)$ and $P_{n+1}(a)$ do not have any common eigenvector associated to 0.*

iii) *If a is a zero of $P_n(x)$ of multiplicity just N, then $P_n(a) = 0_{N\times N}$.*

iv) *If we write $x_{n,k}$, with $k=1,\cdots,nN$, for the zeros of $P_n(x)$ ordered in increasing size (and taking into account their multiplicities), then*

$$x_{n+1,k} \leq x_{n,k} \leq x_{n+1,k+N} \quad \textit{for} \quad k=1,\cdots,nN.$$

Next, A. J. Durán supported in the results of [53] showed two important results. The first is a quadrature formula for matrix polynomials

Theorem 2. ([53]). *Let n be a nonnegative integer. We write $x_{n,k}$, $k=1,\cdots,m$, for the different zeros of the matrix polynomial $P_n(x)$ ordered in increasing size and $\Gamma_{n,k}$ for the matrices*

$$\Gamma_{n,k} = \frac{1}{(\det(P_n(x)))^{(\alpha_k)}(x_{n,k})} (\operatorname{adj}(P_n(t)))^{(\alpha_k - 1)}(x_{n,k}) Q_n(x_{n,k}), \quad k = 1, \cdots, m,$$

where α_k *is the multiplicity of* $x_{n,k}$ *and* $Q_n(x)$ *is the matrix associated polynomial of the first kind. Then for any matrix polynomial* $P(x)$ *with* $\deg(P(x)) < 2n$ *the following formula holds*

$$\int P(x)dW(x) = \sum_{k=1}^{m} P(x_{n,k})\Gamma_{n,k}.$$

Moreover, $\Gamma_{n,k}$ *are positive semidefinite matrices of rank* α_k, $k = 1, \cdots, m$.

Other type of quadrature formulas had already been studied by A. Sinap and W. Van Assche [133] as well as by A. J. Durán and P. López-Rodríguez [61].

The second result that A. J. Durán showed, is the extension of Markov's theorem for matrix orthogonal polynomials, which is a consequence of the above quadrature formula. It states that

$$\lim_{n\to\infty} P_n^{-1}(z)Q_n(z) = \int \frac{dW(x)}{z-x}$$

and the convergence is uniform for compact subsets of $\mathbb{C}\setminus\Gamma$, where $\Gamma = \bigcap_{N\geq 0} \overline{\bigcup_{n\geq N} \sigma(P_n)}$ and $\sigma(P_n)$ is the set of zeros of $P_n(x)$. He also showed that the result is not true for $Q_n(z)P_n^{-1}(z)$.

Since then, the increasing activity in this scientific field has produced a vast bibliography, mainly concerning the extension of results which are known for the scalar case (see for example [35, 53, 61, 122, 133, 134]). However, there are some open problems and unexpected results.

For instance, the classification of sequences of matrix orthogonal polynomials (far away from the classical diagonal cases) satisfying second order linear differential equations with matrix polynomial coefficients which are independent of the degree of the polynomial eigenfunctions, i.e. the so-called Bochner problem, is still open (see [58]). A partial solution to this problem was given by A. J. Durán in [54], where positive definite matrices of measures whose matrix inner product has a symmetric left-hand side matrix second order linear differential operator are characterized. For this purpose, he introduced two matrix second order linear differential operators associated with matrix polynomials $A_2(x), A_1(x), A_0(x)$,

$$\begin{aligned} l_{2,R} &= D''A_2 + D'A_1 + D^0A_0, && \text{right-hand side,} \\ l_{2,L} &= A_2D'' + A_1D' + A_0D^0, && \text{left-hand side.} \end{aligned}$$

In a similar way as in the scalar case, it can be proved that if l_2 (hereinafter l_2 is a left or right hand side matrix second order linear differential operator) is symmetric for $\langle\cdot,\cdot\rangle_W$, then the matrix polynomials $A_2(x)$, $A_1(x)$ and $A_0(x)$ have degree at most 2, 1, and 0, respectively, as well as that the sequence of matrix orthonormal polynomials with respect to $\langle\cdot,\cdot\rangle_W$, satisfies the second order linear differential equation

$$l_2(P_n) = \Gamma_n P_n \qquad n = 0, 1, \ldots \tag{5}$$

for certain Hermitian matrices Γ_n [54]. The converse of the second statement is false for $l_{2,L}$ since non-commutativity of the product of matrices. If the matrix of measures W is considered as a operator acting on $\mathbb{C}^{N\times N}[x]$, that is

$$\langle PW, Q\rangle = \int P(x)dW(x)Q^{\dagger}(x) \quad Q \in \mathbb{C}^{N\times N}[x],$$
$$\langle WP, Q\rangle = \int dW(x)P(x)Q^{\dagger}(x) \quad Q \in \mathbb{C}^{N\times N}[x],$$

and

$$\left\langle W', P\right\rangle = -\left\langle W, P'\right\rangle,$$

then, assuming that $A_0W = WA_0^{\dagger}$ and the operator l_2 is symmetric for $\langle\cdot,\cdot\rangle_W$, we get (see [54])

$$\begin{aligned} (A_2W)' &= A_1W, \\ A_2W &= WA_2^{\dagger}, \\ A_1W &= WA_1^{\dagger}, \end{aligned} \tag{6}$$

and, again, the converse is false for $l_{2,L}$. Note that (6) is Pearson's equation in the scalar case. The above affirmation, allowed to A. J. Durán to prove the following classification theorem

Theorem 3. ([54]). *The following statements are equivalent.*

i) *W is a matrix of measures whose matrix sesquilinear form $\langle\cdot,\cdot\rangle_W$ has a symmetric left-hand side second order linear differential operator.*

ii) *A nonsingular matrix D exists for which $D^{\dagger}XD$, where $X = vI_N$ with v a classical scalar weight (Jacobi, Hermite or Laguerre).*

The above theorem does not hold for a symmetric right-hand side second order linear differential operator (see Appendix in [54]).

Now the question is if there exists a classification theorem for right-hand side second order linear differential equation. This is still an open problem. However a partial result has been obtained.

Theorem 4. ([54]). *Let $W(x)$ be a matrix weight whose sesquilinear form $\langle\cdot,\cdot\rangle_W$ has a right-hand side second order linear differential operator for which A_2 is a nonsingular numerical matrix (not depending on x) and A_0 is the identity matrix up to a multiplicative constant. Then a nonsingular matrix D exists such that $W = D^{\dagger}XD$, where X is a diagonal matrix whose entries in the diagonal are classical Hermite weights up to a linear change of variables.*

From here, a large number of papers has been written in the continuous and discrete case (see [16, 41, 42, 60, 62, 63]). In [58] A. J. Durán and A. Grünbaum gave a survey about techniques for study symmetric right-hand side matrix second order linear differential operator. An important result is the following one.

Theorem 5. ([56, 58]). *Assuming that $dW(x) = W(x)dx$ with a smooth $W(x)$, the following statements are equivalent.*

i) *The operator $l_{2,R}$ is symmetric with respect to $\langle \cdot,\cdot \rangle_W$.*

ii) *The boundary conditions that*

$$A_2(x)W(x) \qquad \textit{and} \qquad (A_2(x)W(x))' - A_1(x)W(x)$$

should have vanishing limits at each of the endpoints of the support of $W(x)$, and the weight matrix $W(x)$ should satisfy

$$A_2W = WA_2^\dagger,$$

$$2(A_2W)' = WA_1^\dagger + A_1W, \tag{7}$$

$$WA_0^\dagger = (A_2W)'' - (A_1W)' + A_0W. \tag{8}$$

If we assume that $WA_1^\dagger = A_1W$, and A_2 is a scalar polynomial, then it implies a scalar type Rodrigues' formula for the sequence of matrix orthogonal polynomials $(P_n(x))_{n\in\mathbb{N}}$ with respect to $\langle \cdot,\cdot \rangle_W$ [57, 58],

$$P_n(x) = (A_2^n(x)W(x))^{(n)}W^{-1}(x).$$

However, with the above hypothesis W reduces to scalar weights. A suitable difference between the scalar and the matrix case is that the equation (7) does not imply that the right-side matrix second differential equation (8), that is, orthogonal polynomials satisfying a "non-commutative Pearson's equation" like (8) do no to satisfy second order linear differential equation like (5) [58]. Indeed,

Theorem 6. ([58]). *If the weight matrix W has a corresponding symmetric second order linear differential operator like (5) with $A_2(x) = a_2(x)I_N$, where $a_2(x)$ is a scalar polynomial, then $\det W(x)$ is a classical scalar weight (up to a scalar change of variable).*

In [57] it was discussed why a second order linear differential equation with coefficients independent of n does not yields in the matrix case a Rodrigues' formula of the type $P_n(x) = C_n(\Phi^n(x)W)^{(n)}W^{-1}$, where Φ is a matrix polynomial of degree not greater than 2 and C_n are nonsingular matrices. Here the role of a scalar type Pearson's equation as well as that of a non-commutative version of it is also mentioned and the following interesting fact is proved.

Theorem 7. ([57]). *Let W be a weight matrix satisfying the matrix Person equation*

$$\big(\phi(x)W'(x)\big) = \Psi(x)W(x)$$

where $\phi(x)$ is a scalar polynomial of degree not greater than 2 and $\Psi(x)$ is a matrix polynomial of degree 1 with nonsingular leading coefficient. We assume that the weight matrix W also satisfies the boundary conditions that $\phi(x)W(x)$ has vanishing limits at each of the end points of the support of $W(x)$. If the degree of $\phi(x)$ is 2 and in addition we assume that its zeros are simple and that the spectrum of the leading coefficient of $\Psi(x)$ is disjoint with the set of natural numbers, then

$$P_n(x) = (\phi^n(x)W(x))^{(n)}W^{-1}$$

is a sequence of matrix polynomials of degree n with not singular coefficients. Moreover they are orthogonal with respect to $W(x)$.

Other important difference between the scalar and the matrix case was showed by M. Castro and A. Grünbaum [31] who find families of matrix orthogonal polynomials satisfying a first order linear differential equation, a fact that does not hold in the scalar case. A example of this fact is also showed in this thesis (see Remark 4.18).

A problem of interest in recent years is the bispectral problem and the generation of new solutions by using the Darboux transformation [78, 79]. In particular, for matrix polynomials satisfying the bispectral problem, new solutions can be generated using Christoffel and Geronimus transformations (see for example [80]). A first step is given in this thesis (see also [11, 12, 68]). We study the perturbation of a matrix of linear functionals consisting in the multiplication by a matrix polynomial of an arbitrary degree such that its leading coefficient is a non-singular matrix (Matrix Christoffel Perturbation). In this way, we also study the matrix analogue of the scalar Geronimus transformation as well as several extensions of them, including left and right multiplication by different matrix polynomials.

Outline of this thesis

This thesis it focused on the study of matrix transformations of matrices with linear functionals as entries. In particular, we study the Christoffel, Geronimus, and Geronimus-Uvarov transformations. In order to this work be self-contained as possible, we have presented it in seven chapters.

- **Chapter 1** contents some preliminary concepts and notation. It is divided in two parts. In the first one, we will set out the basic definitions concerning semi-infinite matrices, modules and matrix polynomials together with their spectral theory. In the second part, we will emphasize the definition of matrix of linear functionals (and their associated sesquilinear form) as well as the concept of bi-orthogonality for matrix polynomials.

- **Chapter 2** is divided in two parts. In the first one, we summarize some results concerning symmetric bilinear forms such that the multiplication operator by a polynomial is symmetric, as well as their relation with recurrence relations of higher order [51]. In the second part we explain carefully the main result of [64] concerning the relation between a scalar sequence of polynomials satisfying a higher order recurrence relation and a sequence of matrix orthonormal polynomials. To conclude this summary, we show some examples of inner products as well as the relation between their corresponding orthogonal polynomials and a sequence of matrix orthogonal polynomials defined below.

- The content of **Chapter 3** is divided in two parts. The first one is based on the works by A. J. Durán [51], A. J. Durán and W. Van Assche [64], and M. Derevyagin and F. Marcellán [48], the latter being the reference work of this first part of chapter. Here, we will give some preliminary results concerning Geronimus spectral transformations in the context of linear functionals and, then, we will motivate the definition of the Geronimus transformation for symmetric bilinear forms defined in terms of a positive measure with the property that the multiplication operator by a polynomial is symmetric with respect to the bilinear form. Next, we will introduce the notion of multiple Geronimus transformation and we generalize the result given in [109], finding, for example, that for certain discrete Sobolev inner product type associated with a multiple Geronimus transformation to a bilinear form yields a simple matrix Geronimus transformation. We have published these results in [47].

  Motivated by the above result, in the second part of this chapter we study the (symmetric) Geronimus transformation but now, on positive definite matrix of measures. i.e, we are interested in the analysis of sesquilinear forms $\langle\cdot,\cdot\rangle_W$ such that

$$\langle P(x)W(x), Q(x)W(x)\rangle_W = \int P(x)dMQ^{\dagger}(x),$$

  where M is a positive definite matrix of measures and $W(x)$ is a fixed matrix polynomial of arbitrary degree. Here we find conditions for the existence of the sequence of matrix orthogonal polynomials with respect to $\langle\cdot,\cdot\rangle_W$ as well as connection formulas between the original and perturbed matrix orthogonal polynomials. We have published these results in [67] and [68].

- **Chapter 4** is devoted to the extension of Christoffel formula for matrix orthogonal polynomials from a more general perspective. More precisely, given a matrix of linear functionals u and $W(x)$ a matrix polynomial with nonsingular leading coefficient, we deal with the following matrix transformation, $\hat{u} = W(x)u$. The main result in this chapter is Theorem 4.11, where we obtain connection formulas between the sequences of original and perturbed matrix monic bi-orthogonal polynomials. For that aim, we use the rich spectral theory available today for this type of polynomials. In particular the Jordan chains and right (left) root polynomials will be extremely useful [77, 110]. Finally, we see that these Christoffel transformations can be extended to more general scenarios in the theory of integrable systems. We have published these results in [11].

- **Chapter 5** is devoted to the extension of the Geronimus transformation for matrices of linear functionals supported in the real line, i.e. we multiply a matrix of linear functionals by the inverse of a matrix polynomial $W_G(x)$ and, then, we add a sum of adequate masses (they depend on the left root polynomials). Here we develop two different methods in order to get the the connection formula between the sequences of original and perturbed matrix bi-orthogonal polynomials: The spectral and non-spectral ones.

- In **Chapter 6** the extension of the Geronimus-Uvarov transformation for matrices of linear functionals supported in the real line is considered. Here, we see this transformation as a composition of a Geronimus transformation and, next, a Christoffel transformation. In terms of matrices of linear functionals will be $u \to \check{u} \to \hat{u}$, where $\check{u}W_G(x) = u$ and $\hat{u} = W_C(x)\check{u}$, with $W_C(x)$ and $W_G(x)$ are matrix polynomials. As in Chapter 5, we obtain the connection formula between the sequences of original and perturbed matrix bi-orthogonal polynomials using again spectral and non-spectral methods. In the spectral method, we find the representation of the perturbed matrix bi-orthogonal polynomials in terms of the family of the original ones and the second kind functions (see (1.9)). Here, we use the fact that the leading coefficients of $W_C(x)$ and $W_G(x)$ are nonsingular matrices. In the non-spectral method, we give a representation of the perturbed bi-orthogonal polynomials without any assumption about the leading coefficient of $W_G(x)$. Finally, as an application, we study the matrix Uvarov transformation, that is, the original functional by adding a finite sum of masses. The results of Chapter 5 and Chapter 6 have been published in [12].

- Finally, in **Chapter 7** a summary of the main results as well as a list of open problems are stated.

Chapter 1

Preliminaries

1.1 Semi-infinite matrices

In this dissertation we will use semi-infinite block matrices (and algebraic operations between them). The shape of these matrices is as follows

$$A=\begin{bmatrix} A_{0,0} & A_{0,1} & \cdots \\ A_{1,0} & A_{1,1} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

In our case, the entries $A_{i,j}$ of the matrix A will be square matrices of size $p\times p$. It is clear that in the semi-infinite case, the product of matrices is not always well defined, for example the product $[1,1/2,1/3\cdots]\begin{bmatrix}1\\1\\1\\\vdots\end{bmatrix}=\sum_{n=1}^{\infty}1/n$ does not converge. Moreover, even if the product between matrices is well defined, the associative law can fail, that is, $(AB)C\neq A(BC)$ for A, B, C block semi-infinite matrices (see [29]). With this in mind, in this section we will give some results related to the manipulation of semi-infinite matrices and, in particular, for Hessenberg type block matrices.

First of all some notation must be needed. For an arbitrary, (finite or infinite) matrix A, the matrices $A^{\top}$ and $A^{\dagger}=\bar{A}^{\top}$ are the transpose and transpose conjugate of A, respectively. Let $\mathbb{C}$ (resp. $\mathbb{R}$) be the set of complex (real) numbers and denote by $\mathbb{C}^{m\times p}$ ($\mathbb{R}^{m\times p}$) the linear space of $m\times p$ matrices with complex (real) entries, In particular $\mathbb{C}^{p}$ and $(\mathbb{C}^{p})^{*}$ denote the spaces of column (or matrices of size $p\times 1$) and row vectors (or matrices of size $1\times p$), respectively. Observe that $(\mathbb{C}^{p})^{*}$ is the dual space of $\mathbb{C}^{p}$.

Definition 1.1. *The product AB of two matrices A and B (non-necessarily square matrices) is said to be admissible if any matrix entry $(AB)_{i,j}=\sum_k A_{i,k}B_{k,j}$ involves only a finite number of non-null terms.*

Notice that the product of finite matrices is always admissible. A special type of matrices are the lower and upper Hessenberg block matrices.

Definition 1.2. *A block matrix A is said to be a lower (resp. upper) block Hessenberg type matrix if there exists $N \in \mathbb{N}$ such that $A_{i,j} = 0$ for $j > i+N$ (resp. $i < N+j$).*

$$\begin{pmatrix} * & * & * & \cdots & * & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & * & \cdots & * & \cdots \\ & * & * & \cdots & * & \cdots \\ & & * & \cdots & * & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Hessenberg upper block matrix shape.

$$\begin{pmatrix} * & * & \cdots & * & & & & & \\ * & * & \cdots & * & * & & & & \\ * & * & \cdots & * & * & * & & & \\ * & * & \cdots & * & * & * & * & & \\ * & * & \cdots & * & * & * & * & * & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Hessenberg lower block matrix shape.

Observe that the product AB is always admissible if A is lower block Hessenberg type or B is an upper Hessenberg block type. Moreover, if A and B are lower (upper) block Hessenberg type matrices, then the product AB is again lower (upper) block Hessenberg type. A special type of matrices are the band block matrices. Band block matrices are precisely those which are simultaneously upper and lower block Hessenberg type. As in the finite case, the product of semi-infinite matrices satisfies the distributive law $A(B+C) = AB+BC$ when the products AB and BC are admissible. Besides, if AB is admissible, then $(AB)^\dagger$ is also admissible and $(AB)^\dagger = B^\dagger A^\dagger$. However, as we stated above, the associative law can fail even if all the involved matrix products are admissible [29].

Proposition 1.3 ([29])**.** *The associative property $(AB)C = A(BC)$ of a matrix product is valid in any of the following cases*

i) *A and B are lower block Hessenberg type.*

ii) *B and C are upper block Hessenberg type.*

iii) *A is lower block Hessenberg type and and B is upper block Hessenberg type.*

In contrast to the finite case, for a semi-infinite block matrix A, A not always has an unique inverse matrix (if there exists one). In other words, the systems $AX = I$ and $YA = I$ can have more than one solution or, X and Y can be unique but $Y \neq X$, even if A is Hessenberg block type [37]. However for this last set of matrices we have the following result.

Proposition 1.4 ([29, 37])**.** *If A is a lower (upper) block Hessenberg type and there exists a matrix B also lower (upper) block Hessenberg type, such that $AB = I$, then B is the unique solution of the systems $AX = I$ and $YA = I$, respectively. In this case B will be denoted by A^{-1}.*

If A is a lower (upper) block Hessenberg type with an unique inverse matrix, then A is said to be an H- matrix.

Corollary 1.5 ([37]). *If A is either a lower or upper triangular block matrix such that the blocks of the main diagonal are nonsingular matrices, then A has an unique inverse.*

In general we will denote by A^{-1} the inverse of the matrix A whenever this inverse is unique. If A and B have an unique inverse, then some rules of manipulation of inverses of finite matrices also apply. For instance, $(A^{\dagger})^{-1} = (A^{-1})^{\dagger}$ will be denoted as $A^{-\dagger}$. In particular, for H matrices $(AB)^{-1} = B^{-1}A^{-1}$ holds.

Remark 1.6. *In this dissertation we always deal with Hessenberg block matrices or matrices that can be factorized in terms of them. Thus, when we need to use the associative law of the product, and the hypothesis of Proposition 1.3 will be satisfied, we will forget the associativity parenthesis.*

Remark 1.7. *With the aim of not saturating the notation, if B is the block semi-infinite matrix*

$$B = \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} & \cdots \\ B_{2,1} & B_{2,2} & & \\ B_{3,1} & & \ddots & \\ \vdots & & & \end{pmatrix}, \tag{1.1}$$

where $B_{i,j}$ is a $p \times p$ matrix, and A is a $p \times p$ matrix, then the product AB will be understood as

$$AB = \begin{pmatrix} A & & \\ & A & \\ & & \ddots \end{pmatrix} \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} & \cdots \\ B_{2,1} & B_{2,2} & & \\ B_{3,1} & & \ddots & \\ \vdots & & & \end{pmatrix} = \begin{pmatrix} AB_{1,1} & AB_{1,2} & AB_{1,3} & \cdots \\ AB_{2,1} & AB_{2,2} & & \\ AB_{3,1} & & \ddots & \\ \vdots & & & \end{pmatrix}.$$

Similarly, BA means

$$BA = \begin{pmatrix} B_{1,1}A & B_{1,2}A & B_{1,3}A & \cdots \\ B_{2,1}A & B_{2,2}A & & \\ B_{3,1}A & & \ddots & \\ \vdots & & & \end{pmatrix}.$$

In particular, if B and C are block semi-infinite Hessenberg matrices with blocks of size $p \times p$, then $B(AC) = (BA)C$ (Proposition 1.3).

1.2 Spectral theory of matrix polynomials

Here some background material concerning the spectral theory of matrix polynomials is introduced. For further reading we refer to [77].

Recall that if R is a ring, then a left module over R is a set M together with two operations

$$+ : M \times M \to M, \text{ and } \cdot : R \times M \to M,$$

such that for $m, n \in M$ and $a, b \in R$ we have

i) $(M, +)$ is an Abelian group.

ii) $(a+b)\cdot m = a\cdot m + b\cdot m$ and $a\cdot(m+n) = a\cdot m + a\cdot n$.

iii) $(a\cdot b)\cdot m = a\cdot(b\cdot m)$.

In a similar way, one defines a right module on R. If M is a left and right module over R, then M is said to be a bi-module [102, 129].
M is said to be a free left (or right) module over R if M admits a basis, that is, there exists a subset S of M such that S is non empty, S generates M, ($M = span(S)$), and S is linearly independent.

Definition 1.8. *Let $A_0, A_1 \cdots, A_N \in \mathbb{C}^{p\times p}$ be square matrices of size $p\times p$ with complex entries. A matrix polynomial $W(x)$ of degree N is a formal expression*

$$W(x) = A_N x^N + A_{N-1}x^{N-1} + \cdots + A_1 x + A_0, \tag{1.2}$$

where we take x as a variable. The matrix polynomial is said to be monic when $A_N = I_p$, where $I_p \in \mathbb{C}^{p\times p}$ denotes the identity matrix. The linear space of matrix polynomials with coefficients in $\mathbb{C}^{p\times p}$ will be denoted by $\mathbb{C}^{p\times p}[x]$.

Remark 1.9. *If $P(x) = x^n + A_{n-1}x^{n-1} + \cdots + A_0$ is a matrix polynomial with $A_i \in \mathbb{C}^{p\times p}$ and B is a block semi-infinite matrix as in (1.1), then we understand $P(B)$ as*

$$P(B) = B^n + A_{n-1}B^{n-1} + \cdots + A_0\mathbb{I}_p,$$

where $\mathbb{I}_p$ is the semi-infinite identity matrix.

Observe that $\mathbb{C}^{p\times p}[x]$ is a free bi-module (and, in particular, a left module) on the ring $\mathbb{C}^{p\times p}$ with basis $\{I_p, xI_p, x^2I_p \ldots\}$. Important submodules of $\mathbb{C}^{p\times p}[x]$ are the sets $\mathbb{C}_n^{p\times p}[x]$ of matrix polynomials of degree less than or equal to n with the basis $\{I_p, xI_p, \ldots x^n I_p\}$ of cardinality $n+1$. Since $\mathbb{C}_n^{p\times p}[x]$ has an invariant basis number [129], then any other basis has the same cardinality. If $(r_n(x))_{n\in\mathbb{N}}$ is a sequence of monic matrix polynomials where each $r_n(x)$ has degree n, then $span(r_n(x))_{n\in\mathbb{N}}$ is a free left module over $\mathbb{C}^{p\times p}$ with basis precisely $(r_n(x))_{n\in\mathbb{N}}$. Notice that $span(r_n(x))_{n\in\mathbb{N}}$ is a submodule of $\mathbb{C}^{p\times p}[x]$. Furthermore, for each $n\in\mathbb{N}$ there exist elements $b_{n,k}\in\mathbb{C}^{p\times p}$, $k = 0,\ldots,n-1$, such that

$$r_n(x) = x^n I_p + \sum_{k=0}^{n-1} b_{n,k}x^k.$$

The above relation reads as

$$\begin{pmatrix} r_0(x) \\ r_1(x) \\ r_2(x) \\ \vdots \end{pmatrix} = L \begin{pmatrix} I_p \\ xI_p \\ x^2 I_p \\ \vdots \end{pmatrix},$$

where L is a semi-infinite lower matrix with 1's as entries in the main diagonal. From the structure of the matrix L we deduce that there exists an unique semi-infinite matrix L^{-1} such that $LL^{-1} = L^{-1}L = \mathbb{I}_p$, (see Corollary 1.5). The above implies that there exists an isomorphism between $\mathbb{C}^{p\times p}[x]$ and $span(r_n(x))_{n\in\mathbb{N}}$ and, therefore, $\mathbb{C}^{p\times p}[x] = span(r_n(x))_{n\in\mathbb{N}}$ and $(r_n(x))_{n\in\mathbb{N}}$ is a basis of $\mathbb{C}^{p\times p}[x]$. In a similar way we get that $\mathbb{C}_n^{p\times p}[x] = span(r_k(x))_{k=0}^n$ for every $n \in \mathbb{N}$.

Definition 1.10. *We say that a matrix polynomial $W(x)$ as in (1.2) is monic normalizable if* $\det A_N \neq 0$ *and, then, $\tilde{W}(x) := A_N^{-1}W(x)$ is its monic normalization.*

Definition 1.11 (Eigenvalues)**.** *The spectrum, or the set of eigenvalues , $\sigma(W)$, of a matrix polynomial $W(x)$ is the zero set of* $\det W(x)$*, i.e.*

$$\sigma(W) := \{\beta \in \mathbb{C} : \det W(\beta) = 0\}.$$

Sometimes we also refer to the set $\sigma(W)$ as the set of zeros of $W(x)$.

Proposition 1.12. *A monic matrix polynomial $W(x)$, with* $\deg W = N$*, has Np (counting multiplicities) eigenvalues or zeros, i.e., we can write*

$$\det W(x) = \prod_{a=1}^{q} (x - x_a)^{\alpha_a},$$

with $Np = \alpha_1 + \cdots + \alpha_q$.

Remark 1.13. *Given the spectrum $\sigma(W) = \{x_1, \dots, x_q\}$, when we need to discuss generic properties associated with an eigenvalue, and there is no needed to specify which, for the sake of simplicity will denote such an eigenvalue by x_a. Thus, x_a could be any of the eigenvalues x_1, $x_2, \dots, x_q$.*

Proposition 1.14. *The spectrum, eigenvalues, and multiplicities of $W(x)$ and of $\big(W(x)\big)^\top$ coincide. Consequently,*

$$\dim \operatorname{Ker}\Big(\big(W(x)\big)^\top\Big) = \dim \operatorname{Ker}(W(x)).$$

Proof. It follows from the fact that $\det\Big(\big(W(x)\big)^\top\Big) = \det W(x)$. ∎

Definition 1.15 (Linearization [77])**.**

i) *Two matrix polynomials $W_1, W_2 \in \mathbb{C}^{m\times m}[x]$ are said to be equivalent $W_1 \sim W_2$ if there exist two matrix polynomials $E, F \in \mathbb{C}^{m\times m}[x]$, with constant determinants (not depending on x), such that $W_1(x) = E(x)W_2(x)F(x)$.*

ii) *A degree one matrix polynomial $I_{Np}x - A \in \mathbb{C}^{Np\times Np}[x]$ is called a linearization of a monic matrix polynomial $W \in \mathbb{C}^{p\times p}[x]$ if*

$$I_{Np}x - A \sim \begin{bmatrix} W(x) & 0_{p\times(N-1)p} \\ 0_{(N-1)p\times p} & I_{(N-1)p} \end{bmatrix}.$$

Definition 1.16 (Companion matrix [77])**.** *Given a matrix polynomial $W(x) = I_p x^N + A_{N-1}x^{N-1} + \cdots + A_0$ its companion matrix $C_1 \in \mathbb{C}^{Np\times Np}$ is*

$$C_1 := \begin{bmatrix} 0_p & I_p & 0_p & \dots & 0_p \\ 0_p & 0_p & I_p & \ddots & 0_p \\ \vdots & \vdots & \ddots & \ddots & \\ 0_p & 0_p & 0_p & & I_p \\ -A_0 & -A_1 & -A_2 & \dots & -A_{N-1} \end{bmatrix}.$$

The companion matrix plays an important role in the study of the spectral properties of a matrix polynomial $W(x)$, see for example [77, 110] and [111].

Proposition 1.17 ([77])**.** *Given a monic matrix polynomial $W(x) = I_p x^N + A_{N-1}x^{N-1} + \cdots + A_0$ its companion matrix C_1 provides a linearization*

$$I_{Np}x - C_1 = E(x) \begin{bmatrix} W(x) & 0_{p\times(N-1)p} \\ 0_{(N-1)p\times p} & I_{(N-1)p} \end{bmatrix} F(x),$$

where

$$E(x) = \begin{bmatrix} B_{N-1}(x) & B_{N-2}(x) & B_{N-3}(x) & \dots & B_1(x) & B_0(x) \\ -I_p & 0_p & 0_p & \dots & 0_p & 0_p \\ 0_p & -I_p & 0_p & \dots & 0_p & 0_p \\ 0_p & 0_p & -I_p & & 0_p & 0_p \\ \vdots & & & \ddots & \vdots & \\ 0_p & 0_p & 0_p & & -I_p & 0_p \end{bmatrix},$$

$$F(x) = \begin{bmatrix} I_p & 0_p & 0_p & \dots & 0_p & 0_p \\ -I_p x & I_p & 0_p & \dots & 0_p & 0_p \\ 0_p & -I_p x & I_p & \ddots & 0_p & 0_p \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0_p & 0_p & 0_p & \ddots & I_p & 0_p \\ 0_p & 0_p & 0_p & & -I_p x & I_p \end{bmatrix},$$

with $B_{r+1}(x) = xB_r(x) + A_{N-r-1}$ and the initial condition $B_0(x) := I_p$ for $r \in \{0,1,\dots,N-2\}$.

From here one deduces

Proposition 1.18 ([77]). *The eigenvalues, with multiplicities, of a monic matrix polynomial coincide with those of its companion matrix.*

Proposition 1.19 ([77]). *Any nonsingular matrix polynomial $W(x) \in \mathbb{C}^{p\times p}[x]$, $\det W(x) \neq 0$, can be represented*

$$W(x) = E_{x_a}(x)\,\mathrm{diag}((x-x_a)^{\kappa_1},\dots,(x-x_a)^{\kappa_m})F_{x_a}(x)$$

at $x = x_a \in \mathbb{C}$, where $E_{x_a}(x)$ and $F_{x_a}(x)$ are nonsingular matrices and $\kappa_1 \leq \cdots \leq \kappa_m$ are nonnegative integers. Moreover, $\{\kappa_1,\dots,\kappa_m\}$ are uniquely determined by W and they are known as partial multiplicities of $W(x)$ at x_a.

Definition 1.20. *Let x_a be a eigenvalue of a monic matrix polynomial $W(x) \in \mathbb{C}^{p\times p}[x]$. Then*

i) *A non-zero vector $r_0 \in \mathbb{C}^p$ is said to be a right eigenvector, with eigenvalue $x_a \in \sigma(W(x))$, whenever $W(x_a)r_0$ gives the zero vector in $\mathbb{C}^p$, $W(x_a)r_0 = 0$, i.e., $r_0 \in \operatorname{Ker} W(x_a) \neq \{0\}$.*

ii) *A non-zero covector $l_0 \in \left(\mathbb{C}^p\right)^*$ is said to be a left eigenvector, with eigenvalue $x_a \in \sigma(W(x))$, whenever $l_0W(x_a)$, is the zero covector in $\mathbb{C}^p$, i.e., $l_0W(x_a) = 0, \left(l_0\right)^\top \in \left(\operatorname{Ker}(W(x_a))\right)^\perp = \operatorname{Ker}\left((W(x_a))^\top\right) \neq \{0\}$.*

iii) *A sequence of vectors $\{r_0, r_1,\dots, r_{m-1}\}$ is said to be a right Jordan chain of length m corresponding to the eigenvalue $x_a \in \sigma(W)$, if r_0 is an right eigenvector of $W(x_a)$ and*

$$\sum_{s=0}^{j} \frac{1}{s!}\frac{\mathrm{d}^s W}{\mathrm{d}x^s}\Big|_{x=x_a} r_{j-s} = 0, \qquad\qquad j \in \{0,\dots,m-1\}.$$

iv) *A sequence of covectors $\{l_0, l_1 \dots, l_{m-1}\}$ is said to be a left Jordan chain of length m, corresponding to $x_a \in \sigma(W^\top)$, if $\{(l_0)^\top, (l_1)^\top,\dots,(l_{m-1})^\top\}$ is a right Jordan chain of length m for the matrix polynomial $\left(W(x)\right)^\top$.*

v) *A right root polynomial at x_a is a non-zero vector polynomial $r(x) \in \mathbb{C}^p[x]$ such that $W(x)r(x)$ has a zero of certain order at $x = x_a$. The order of this zero is called the order of the root polynomial. Analogously, a left root polynomial is a non-zero covector polynomial $l(x) \in (\mathbb{C}^p)^*[x]$ such that $l(x_a)W(x_a) = 0$.*

vi) *The maximal lengths, either of right or left Jordan chains corresponding to the eigenvalue x_a, are called the multiplicity of the eigenvector r_0 or l_0. They will be denoted by $m(r_0)$ or $m(l_0)$, respectively.*

Proposition 1.21. *Given an eigenvalue $x_a \in \sigma(W(x))$ of a monic matrix polynomial $W(x)$, multiplicities of right and left eigenvectors coincide and are equal to the corresponding partial multiplicities κ_i.*

The above definition generalizes the concept of Jordan chain for matrix polynomials of degree one.

Proposition 1.22. *The Taylor expansion of a right root polynomial $r(x)$ (respectively, of a left root polynomial $l(x)$) at a given eigenvalue $x_a \in \sigma(W)$ of a monic matrix polynomial $W(x)$,*

$$r(x) = \sum_{j=0}^{\kappa-1} r_j (x-x_a)^j, \qquad \text{respectively } l(x) = \sum_{j=0}^{\kappa-1} l_j (x-x_a)^j,$$

provides us with the right Jordan chain

$$\{r_0, r_1, \dots, r_{\kappa-1}\}, \qquad \text{respectively, left Jordan chain } \{l_0, l_1, \dots, l_{\kappa-1}\}.$$

Proposition 1.23 ([77, 110])**.** *Given an eigenvalue $x_a \in \sigma(W)$ of a monic matrix polynomial $W(x)$, with multiplicity $s = \dim \operatorname{Ker} W(x_a)$, we can construct s right root polynomials, (respectively, left root polynomials), for $i \in \{1, \dots, s\}$,*

$$r_i(x) = \sum_{j=0}^{\kappa_i-1} r_{i,j} (x-x_a)^j, \qquad (\text{respectively } l_i(x) = \sum_{j=0}^{\kappa_i-1} l_{i,j} (x-x_a)^j),$$

where $r_i(x)$ are right root polynomials (respectively, $l_i(x)$ are left root polynomials) with the largest order κ_i among all right root polynomials, whose right eigenvector does not belong to $\mathbb{C}\{r_{0,1}, \dots, r_{0,i-1}\}$, (respectively left root polynomials whose left eigenvector does not belong to $\mathbb{C}\{l_{0,1}, \dots, l_{0,i-1}\}$).

Definition 1.24 (Canonical Jordan chains [77, 110])**.** *A canonical set of right Jordan chains (respectively, left Jordan chains) of the monic matrix polynomial $W(x)$ corresponding to the eigenvalue $x_a \in \sigma(W)$ is, in terms of the right root polynomials, (respectively, left root polynomials) described in Proposition 1.23, the following sets of vectors*

$$\{r_{1,0} \dots, r_{1,\kappa_1-1}, \dots, r_{s,0} \dots, r_{s,\kappa_r-1}\}, \quad (\text{respectively, covectors } \{l_{1,0} \dots, l_{1,\kappa_1-1}, \dots, l_{s,0} \dots, l_{s,\kappa_r-1}\}).$$

Proposition 1.25. *For a monic matrix polynomial $W(x)$, the lengths $\{\kappa_1, \dots, \kappa_r\}$ of the Jordan chains in a canonical set of Jordan chains of $W(x)$ corresponding to the eigenvalue x_a, (see Definition 1.24), are the nonzero partial multiplicities of $W(x)$ at $x = x_a$ described in Proposition 1.19.*

Definition 1.26 (Canonical Jordan chains and root polynomials [77, 110])**.** *For each eigenvalue $x_a \in \sigma(W)$ of a monic matrix polynomial $W(x)$, with multiplicity α_a and $s_a = \dim \operatorname{Ker} W(x_a)$, $a \in \{1, \dots, q\}$, we choose a canonical set of right Jordan chains, (respectively, left Jordan chains)*

$$\left\{r^{(a)}_{j,0}, \dots, r^{(a)}_{j,\kappa^{(a)}_j-1}\right\}_{j=1}^{s_a}, \qquad \left(\text{respectively, } \left\{l^{(a)}_{j,0}, \dots, l^{(a)}_{j,\kappa^{(a)}_j-1}\right\}_{j=1}^{s_a}\right),$$

and, consequently, with partial multiplicities satisfying $\sum_{j=1}^{s_a} \kappa^{(a)}_j = \alpha_a$. Thus, we can consider the following adapted right root polynomials

$$r^{(a)}_j(x) = \sum_{l=0}^{\kappa^{(a)}_j-1} r^{(a)}_{j,l} (x-x_a)^l, \quad (\text{respectively left root polynomials } l^{(a)}_j(x) = \sum_{l=0}^{\kappa^{(a)}_j-1} l^{(a)}_{j,l} (x-x_a)^l). \tag{1.3}$$

Definition 1.27 (canonical Jordan pairs [77])**.** *We also define the corresponding canonical Jordan pair (X_a, J_a), where X_a is the matrix*

$$X_a := \left[r^{(a)}_{1,0}, \dots, r^{(a)}_{1,\kappa^{(a)}_1-1}, \dots, r^{(a)}_{s_a,0}, \dots, r^{(a)}_{s_a,\kappa^{(a)}_{s_a}-1}\right] \in \mathbb{C}^{p\times\alpha_a},$$

and J_a is the matrix

$$J_a := \operatorname{diag}(J_{a,1}, \dots, J_{a,s_a}) \in \mathbb{C}^{\alpha_a\times\alpha_a}.$$

Here $J_{a,j} \in \mathbb{C}^{\kappa^{(a)}_j\times\kappa^{(a)}_j}$ are the Jordan blocks of the eigenvalue $x_a \in \sigma(W)$. Then, we say that (X,J) with

$$X := \left[X_1, \dots, X_q\right] \in \mathbb{C}^{p\times Np}, \qquad J := \operatorname{diag}(J_1, \dots, J_q) \in \mathbb{C}^{Np\times Np}, \tag{1.4}$$

is a canonical Jordan pair for $W(x)$.

We have the important result, see [77],

Proposition 1.28. *The Jordan pairs (X_a, J_a) of a monic matrix polynomial $W(x)$ satisfy*

$$A_0X_a + A_1X_aJ_a + \cdots + A_{N-1}X_a(J_a)^{N-1} + X_a(J_a)^N = 0_{p\times\alpha_a},$$

and

$$A_0X + A_1XJ + \cdots + A_{N-1}XJ^{N-1} + XJ^N = 0_{p\times Np},$$

where X and J are defined in (1.4). A key property, see Theorem 1.20 of [77], is

Proposition 1.29. *For any Jordan pair (X,J) of a monic matrix polynomial $W(x) = I_px^N + A_{N-1}x^{N-1} + \cdots + A_0$ the matrix*

$$\begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{N-1} \end{bmatrix} \in \mathbb{C}^{Np\times Np}$$

is nonsingular.

Moreover, Theorem 1.23 of [77] gives the following characterization

Proposition 1.30. *Two matrices $X \in \mathbb{C}^{p\times Np}$ and $J \in \mathbb{C}^{Np\times Np}$ constitute a Jordan pair of a monic matrix polynomial $W(x) = I_px^N + A_{N-1}x^{N-1} + \cdots + A_0$ if and only if the two following properties hold*

i) *The matrix*

$$\begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{N-1} \end{bmatrix}$$

is nonsingular.

ii)

$$A_0X + A_1XJ + \cdots + A_{N-1}XJ^{N-1} + XJ^N = 0_{p\times Np}.$$

Definition 1.31 ([77, 110])**.** *Let $W(x) = I_p x^N + \sum_{k=0}^{N} A_k x^k$ be a matrix polynomial of degree N. We say that $W_1(x)$ with degree m is a right divisor of $W(x)$ if there exists a matrix polynomial $R(x)$ with degree $N-m$ such that*

$$W(x) = R(x)W_1(x).$$

Proposition 1.32 (Corollary 3.8, [77])**.** *A polynomial $W_1(x)$ with degree m is said to be a right divisor of $W(x) = I_p x^N + \sum_{k=0}^{N} A_k x^k$ if and only if for a Jordan pair (X_1, J_1) of $W_1(x)$,*

$$A_0X_1 + \cdots A_{N-1}X_1J_1^{N-1} + X_1J_1^N = 0_{p\times mp},$$

holds.

Proposition 1.33. *Given a monic matrix polynomial $W(x)$, the adapted root polynomials given in Definition 1.26 satisfy*

$$\big(W(x)r_j^{(a)}(x)\big)_{x=x_a}^{(m)} = 0, \qquad \big(l_j^{(a)}(x)W(x)\big)_{x=x_a}^{(m)} = 0, \qquad m \in \{0,\dots,\kappa_j^{(a)}-1\}, \qquad j \in \{1\dots,s_a\}.$$

Here, for a function $f(x)$ we use the following notation for its derivatives evaluated at an eigenvalue $x_a \in \sigma(W)$

$$(f(x))_{x=x_a}^{(m)} := \left.\frac{\mathrm{d}^m f(x)}{\mathrm{d}x^m}\right|_{x=x_a}.$$

In this dissertation we assume that the partial multiplicities are ordered in an increasing way, i.e., $\kappa_1^{(a)} \leq \kappa_2^{(a)} \leq \cdots \leq \kappa_{s_a}^{(a)}$.

Proposition 1.34. *If $r_i^{(a)}(x)$ and $l_j^{(a)}(x)$ are right and left root polynomials corresponding to the eigenvalue $x_a \in \sigma(W(x))$, then a polynomial*

$$w_{i,j}^{(a)}(x) = \sum_{m=0}^{d_{i,j}^{(a)}} w_{i,j;m}^{(a)} x^m \in \mathbb{C}[x], \qquad\qquad d_{i,j}^{(a)} := \kappa_{\min(i,j)}^{(a)} + N - 2,$$

exists such that

$$l_i^{(a)}(x)W(x)r_j^{(a)}(x) = (x-x_a)^{\kappa_{\max(i,j)}^{(a)}} w_{i,j}^{(a)}(x).$$

Proof. From Proposition 1.33 it follows that there exist a covector polynomial $T_1(x)$ and a vector polynomial $T_2(x)$, both of degree N, such that

$$l_i^{(a)}(x)W(x) = (x-x_a)^{\kappa_i^{(a)}} T_1(x), \qquad W(x)r_j^{(a)}(x) = (x-x_a)^{\kappa_j^{(a)}} T_2(x). \tag{1.5}$$

Thus

$$l_i^{(a)}(x)W(x)r_j^{(a)}(x) = (x-x_a)^{\kappa_i^{(a)}} T_1(x)r_j^{(a)}(x), \quad l_i^{(a)}(x)W(x)r_j^{(a)}(x) = (x-x_a)^{\kappa_j^{(a)}} l_i^{(a)}(x)T_2(x),$$

and the result follows. ∎

Definition 1.35 (Spectral jets)**.** *Given a matrix function $f(x)$ which is smooth in its domain of definition, we consider its matrix spectral jets*

$$\begin{aligned}
\mathtt{J}_f^{(j)}(x_a) &:= \left[f(x_a), \dots, \frac{f^{(\kappa_j^{(a)}-1)}(x_a)}{(\kappa_j^{(a)}-1)!} \right] \in \mathbb{C}^{p\times p\kappa_j^{(a)}},\\
\mathtt{J}_f(x_a) &:= \left[\mathtt{J}_f^{(1)}(x_a), \dots, \mathtt{J}_f^{(s_a)}(x_a) \right] \in \mathbb{C}^{p\times p\alpha_a},\\
\mathtt{J}_f &:= \left[\mathtt{J}_f(x_1), \dots, \mathtt{J}_f(x_q) \right] \in \mathbb{C}^{p\times Np^2},
\end{aligned}$$

and given a Jordan pair the root spectral jet vectors

$$\begin{aligned}
\mathcal{J}_f^{(j)}(x_a) &:= \left[f(x_a)r_j^{(a)}(x_a), \dots, \frac{(f(x)r_j^{(a)}(x))_{x=x_a}^{(\kappa_j^{(a)}-1)}}{(\kappa_j^{(a)}-1)!} \right] \in \mathbb{C}^{p\times \kappa_j^{(a)}},\\
\mathcal{J}_f(x_a) &:= \left[\mathcal{J}_f^{(1)}(x_a) \dots, \mathcal{J}_f^{(s_a)}(x_a) \right] \in \mathbb{C}^{p\times \alpha_a},\\
\mathcal{J}_f &:= \left[\mathcal{J}_f(x_1), \dots, \mathcal{J}_f(x_q) \right] \in \mathbb{C}^{p\times Np}.
\end{aligned}$$

Definition 1.36. *We consider the following jet matrices*

$$\begin{aligned}
\mathcal{Q}_{n;j}^{(a)} &:= \mathcal{J}_{I_p x^n}^{(j)}(x_a) = \left[(x_a)^n r_j^{(a)}(x_a), \left(x^n r_j^{(a)}(x)\right)_{x=x_a}^{(1)}, \dots, \frac{\left(x^n r_j^{(a)}(x)\right)_{x=x_a}^{(\kappa_j^{(a)}-1)}}{(\kappa_j^{(a)}-1)!} \right] \in \mathbb{C}^{p\times \kappa_j^{(a)}},\\
\mathcal{Q}_n^{(a)} &:= \mathcal{J}_{I_p x^n}(x_a) = \left[\mathcal{Q}_{n;1}^{(a)}, \dots, \mathcal{Q}_{n;s_a}^{(a)} \right] \in \mathbb{C}^{p\times \alpha_a},\\
\mathcal{Q}_n &:= \mathcal{J}_{I_p x^n} = \left[\mathcal{Q}_n^{(1)}, \dots, \mathcal{Q}_n^{(q)} \right] \in \mathbb{C}^{p\times Np},\\
\mathcal{Q} &:= \mathcal{J}_{\chi_{[N]}} = \begin{bmatrix} \mathcal{Q}_0 \\ \vdots \\ \mathcal{Q}_{N-1} \end{bmatrix} \in \mathbb{C}^{Np\times Np},
\end{aligned}$$

where $(\chi_{[N]}(x))^\top := \left[I_p, \dots, I_p x^{N-1} \right] \in \mathbb{C}^{p\times Np}[x]$.

Lemma 1.37 (Root spectral jets and Jordan pairs)**.** *Given a canonical Jordan pair (X,J), for the monic matrix polynomial $W(x)$ we have*

$$Q_n = XJ^n, \qquad n\in\mathbb{N}.$$

Thus, any polynomial $P_n(x)=\sum_{j=0}^{n} B_j x^j$ has as its spectral jet vector corresponding to $W(x)$ the following matrix

$$\mathcal{J}_{P_n} = B_0X + B_1XJ + \cdots + B_nXJ^n.$$

Proof. The computation

$$\begin{aligned}
\frac{1}{m!}\big(x^n r_j^{(a)}(x)\big)^{(m)}_{x=x_a} &= \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}(x^n)^{(m-k)}_{x=x_a}\big(r_j^{(a)}(x)\big)^{(k)}_{x=x_a}\\
&= \frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}\frac{n!}{(n-m+k)!}(x_a)^{n-m+k}k!r^{(a)}_{j,k}\\
&= \sum_{k=0}^{m}\binom{n}{m-k}(x_a)^{n-m+k}r^{(a)}_{j,k}\\
&= \begin{bmatrix} r^{(a)}_{j,0},\dots,r^{(a)}_{j,m}\end{bmatrix}\begin{bmatrix}(x_a)^{n-m}\binom{n}{m}\\ \vdots \\ x_a^n\binom{n}{n}\end{bmatrix}
\end{aligned}$$

leads

$$\begin{aligned}
Q^{(a)}_{n;j} &= \begin{bmatrix} r^{(a)}_{j,0},\dots,r^{(a)}_{j,\kappa_j^{(a)}-1}\end{bmatrix}\begin{bmatrix} x_a^n & x_a^{n-1}\binom{n}{1} & \cdots & x_a^{n-\kappa_j^{(a)}+1}\binom{n}{\kappa_j^{(a)}-1}\\ 0 & x_a^n & \cdots & x_a^{n-\kappa_j^{(a)}+2}\binom{n}{\kappa_j^{(a)}-2}\\ \vdots & \ddots & \ddots & \vdots\\ 0 & \dots & & x_a^n\end{bmatrix}\\
&= \begin{bmatrix} r^{(a)}_{j,0},\dots,r^{(a)}_{j,\kappa_j^{(a)}-1}\end{bmatrix}(J_{a,j})^n .
\end{aligned}$$

Consequently, in terms of the Jordan pairs associated with the right root polynomials, we have $X_a(J_a)^n = Q_n^{(a)}$ and

$$Q_n = XJ^n.$$

Moreover the matrix Q is nonsingular, see Propositions 1.29 and 1.30. ∎

Definition 1.38. *If $W(x)=\sum\limits_{k=0}^{N}A_kx^k\in\mathbb{C}^{p\times p}[x]$ is a matrix polynomial of degree N, we introduce the matrix*

$$\mathcal{B}:=\begin{bmatrix} A_1 & A_2 & A_3 & \dots & A_{N-1} & A_N\\ A_2 & A_3 & \vdots & \iddots & A_N & 0_p\\ A_3 & \dots & A_{N-1} & \iddots & 0_p & 0_p\\ \vdots & \iddots & \iddots & \iddots & & \vdots\\ A_{N-1} & A_N & 0_p & & & \\ A_N & 0_p & 0_p & & \dots & 0_p\end{bmatrix}\in\mathbb{C}^{Np\times Np}.$$

Definition 1.39 (Jordan triple)**.** *Given*

$$Y=\begin{bmatrix}Y_1\\ \vdots\\ Y_q\end{bmatrix}\in\mathbb{C}^{Np\times p},$$

with $Y_a\in\mathbb{C}^{\alpha_a\times p}$, (X,J,Y) said to be a Jordan triple whenever

$$\begin{bmatrix}X\\ XJ\\ \vdots\\ XJ^{N-1}\end{bmatrix}Y=\begin{bmatrix}0_p\\ \vdots\\ 0_p\\ I_p\end{bmatrix}.$$

Lemma 1.40. *Given a Jordan triple (X,J,Y), for the monic matrix polynomial $W(x)$ we have*

$$\mathcal{Q}=\begin{bmatrix}X\\ XJ\\ \vdots\\ XJ^{N-1}\end{bmatrix}, \qquad\qquad (\mathcal{B}\mathcal{Q})^{-1}=\begin{bmatrix}Y,JY,\dots,J^{N-1}Y\end{bmatrix}=:\mathcal{R}.$$

Proof. From Lemma 1.37 we deduce that

$$\mathcal{Q}=\begin{bmatrix}X\\ XJ\\ \vdots\\ XJ^{N-1}\end{bmatrix},$$

which is a nonsingular matrix, (see Propositions 1.29 and 1.30). The bi-orthogonality condition (2.6) of [77] for $\mathcal{R}$ and $\mathcal{Q}$ is

$$\mathcal{R}\mathcal{B}\mathcal{Q}=I_{Np},$$

and if (X,J,Y) is a canonical Jordan triple it is also shown that

$$\mathcal{R} = \left[Y, JY, \ldots, J^{N-1}Y\right].$$

■

Proposition 1.41. *The matrix* $\mathcal{R}_n := \left[Y, JY, \ldots, J^{n-1}Y\right] \in \mathbb{C}^{Np\times np}$ *has full rank.*

On the other hand,

Definition 1.42. *Let us introduce the bivariate matrix polynomial*

$$\mathcal{V}(x,y) := \left((\chi(y))_{[N]}\right)^{\top} \mathcal{B}(\chi(x))_{[N]} \in \mathbb{C}^{p\times p}[x,y].$$

We consider the complete homogeneous symmetric polynomials in two variables and total degree n.

$$h_n(x,y) = \sum_{j=0}^{n} x^j y^{n-j}.$$

For example,

$$h_0(x,y) = 1, \quad h_1(x,y) = x+y, \quad h_2(x,y) = x^2+xy+y^2, \quad h_3(x,y) = x^3+x^2y+xy^2+y^3.$$

Proposition 1.43. *In terms of the complete homogeneous symmetric polynomials in two variables we can write*

$$\mathcal{V}(x,y) = \sum_{j=1}^{N} A_j h_{j-1}(x,y).$$

1.3 Sesquilinear forms and orthogonal matrix polynomials

Recall that the polynomial ring $\mathbb{C}^{p\times p}[x]$ is a free bimodule over the ring of matrices $\mathbb{C}^{p\times p}$ with a basis given by $\{I_p, I_p x, I_p x^2, \ldots\}$. Lets us also recall that for the bi-submodule (which is free again) $\mathbb{C}_m^{p\times p}[x]$, any basis has cardinality $m+1$.

Definition 1.44 (Sesquilinear form)**.** *A sesquilinear form* $\langle\cdot,\cdot\rangle$ *on the bimodule* $\mathbb{C}^{p\times p}[x]$ *(with real variable) is a map*

$$\langle\cdot,\cdot\rangle : \mathbb{C}^{p\times p}[x] \times \mathbb{C}^{p\times p}[x] \longrightarrow \mathbb{C}^{p\times p},$$

such that for any triple $P,Q,R \in \mathbb{C}^{p\times p}[x]$ *of matrix polynomials we have*

i) $\langle AP(x)+BQ(x), R(x)\rangle = A\langle P(x),R(x)\rangle + B\langle Q(x),R(x)\rangle$, *for all* $A,B \in \mathbb{C}^{p\times p}$.

ii) $\langle P(x), AQ(x)+BR(x)\rangle = \langle P(x),Q(x)\rangle A^{\dagger} + \langle P(x),R(x)\rangle B^{\dagger}$, *for all* $A,B \in \mathbb{C}^{p\times p}$.

If $\langle P(t), Q(t)\rangle = \langle Q(t), P(t)\rangle^{\dagger}$, then $\langle\cdot,\cdot\rangle$ is called a Hermitian sesquilinear form.

Definition 1.45. *A bilinear form $B(\cdot,\cdot)$ defined on the set of polynomials with real coefficients $\mathbb{R}[x]$ is a mapping*

$$B : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$$

that satisfies

i) $B(p+h,q) = B(p,q) + B(h,q)$ *and* $B(p,h+q) = B(p,h) + B(p,q)$,

ii) $\alpha B(p,q) = B(\alpha p,q) = B(p,\alpha q)$,

where p, q, h belong to $\mathbb{R}[x]$ and α is a real number. $B(\cdot,\cdot)$ is said to be a symmetric bilinear form if $B(f,g) = B(g,f)$ for every $f,g \in \mathbb{R}[x]$.

For any pair of matrix polynomials $P = \sum_{k=0}^{\deg P} p_k x^k$ and $Q(x) = \sum_{l=0}^{\deg Q} q_l x^l$, the sesquilinear form is defined by

$$\langle P(x), Q(x)\rangle = \sum_{\substack{k=1,\dots,\deg P\\ l=1,\dots,\deg Q}} p_k m_{k,l} (q_l)^{\dagger}.$$

Here the coefficients are the values of the sesquilinear form on the basis of the module, i.e. $m_{k,l} = \left\langle x^k 1_p, x^l 1_p\right\rangle$. The semi-infinite matrix

$$M = \begin{bmatrix} m_{0,0} & m_{0,1} & \cdots \\ m_{1,0} & m_{1,1} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

is known as the block matrix of moments (or Gram matrix) of the sesquilinear form. The k-th truncation of M is denoted by

$$M_{[k]} := \begin{bmatrix} m_{0,0} & \cdots & m_{0,k-1} \\ \vdots & & \vdots \\ m_{k-1,0} & \cdots & m_{k-1,k-1} \end{bmatrix}.$$

Definition 1.46 (Bi-orthogonal matrix polynomials)**.** *Given a sesquilinear form $\langle\cdot,\cdot\rangle$, two sequences of polynomials $(P_n^{[1]}(x))_{n\in\mathbb{N}}$ and $\left(P_n^{[2]}(x)\right)_{n\in\mathbb{N}}$ are said to be bi-orthogonal with respect to $\langle\cdot,\cdot\rangle$ if*

i) $\deg(P_n^{[1]}) = \deg(P_n^{[2]}) = n,$ *for all $n \in \mathbb{N}$.*

ii) $\left\langle P_n^{[1]}(x), P_m^{[2]}(x)\right\rangle = \delta_{n,m} H_n,$ *for all $n,m \in \mathbb{N}$,*

where H_n are nonsingular matrices and $\delta_{n,m}$ is the Kronecker delta.

Here, it is important to notice the order of the polynomials in the sesquilinear form; i.e., if $n \neq m$, then $\left\langle P_n^{[2]}(x), P_m^{[1]}(x) \right\rangle$ could be different from $0_{p\times p}$.

Definition 1.47. i) *The sesquilinear form $\langle\cdot,\cdot\rangle$ is said to be quasi-definite if the leading principal sub-matrices of the corresponding block matrix of moments are nonsingular (i.e, it is quasi-definite). It is said to be positive definite if $\langle\cdot,\cdot\rangle$ is Hermitian and $\langle R(t), R(t)\rangle$ is a positive definite matrix for all $R(t) \in \mathbb{C}^{p\times p}[x]$ with nonsingular leading coefficient. Indeed, any quasi-definite sesquilinear form will have sequences of matrix monic bi-orthogonal polynomials [19].*

ii) *A symmetric bilinear form is said to be quasi-definite (resp. positive definite) if all leading principal submatrices of its Gram matrix are nonsingular (positive definite).*

Examples of sesquilinear forms in $\mathbb{C}^{p\times p}[x]$ can be given by matrices with complex (or real) measures as entries

$$\mu = \begin{bmatrix} \mu_{1,1} & \cdots & \mu_{1,p} \\ \vdots & & \vdots \\ \mu_{p,1} & \cdots & \mu_{p,p} \end{bmatrix},$$

i.e., a $p \times p$ matrix of Borel measures in $\mathbb{R}$. Given any pair of matrix polynomials $P(x), Q(x) \in \mathbb{C}^{p\times p}[x]$ we introduce the following sesquilinear form

$$\langle P(x), Q(x)\rangle_\mu = \int_{\mathbb{R}} P(x)\,\mathrm{d}\mu(x)(Q(x))^\dagger.$$

A more general sesquilinear form can be constructed in terms of linear functionals or generalized functions. In [113, 114] a linear functional setting for orthogonal polynomials is given. We consider the space of polynomials $\mathbb{C}[x]$, with an appropriate topology, as the space of fundamental solutions, in the sense of [74, 75], and take the space of linear functionals as the corresponding continuous linear functionals. It is remarkable that the topological dual space coincides with the algebraic dual space. On the other hand, this space of linear functionals is the space of formal series with complex coefficients $(\mathbb{C}[x])' = \mathbb{C}[\![x]\!]$.
In Chapter 4 we use linear functionals with a well defined support and, consequently, the previously described setting requires of a suitable modification. Following [131, 74, 75], let us recall that the space of distributions is a space of linear functionals when the space of fundamental functions is the complex valued smooth functions of compact support $\mathcal{D} := C_0^\infty(\mathbb{R})$, the so called space of test functions. In this context, the set of zeros of a distribution $u \in \mathcal{D}'$is the region $\Omega \subset \mathbb{R}$ if for any fundamental function $f(x)$ with support in Ω we have $\langle u, f\rangle = 0$. Its complement, a closed set, is called support, $\operatorname{supp} u$, of the distribution u. Distributions of compact support, $u \in \mathcal{E}'$, are linear functionals for which the space of fundamental functions is the topological space of complex valued smooth functions $\mathcal{E} = C^\infty(\mathbb{R})$. As $\mathbb{C}[x] \subsetneq \mathcal{E}$ we also know that $\mathcal{E}' \subsetneq (\mathbb{C}[x])' \cap \mathcal{D}'$. The set of distributions of compact support is a first example of an appropriate framework for the consideration of polynomials and supports simultaneously. More general settings appear within the

space of tempered distributions $\mathcal{S}'$, $\mathcal{S}' \subsetneq \mathcal{D}'$. The space of fundamental functions is given by the Schwartz space $\mathcal{S}$ of complex valued fast decreasing functions, see [74, 75, 131]. We consider the space of fundamental functions constituted by smooth functions of slow growth $\mathcal{O}_M \subset \mathcal{E}$, whose elements are smooth functions with derivatives bounded by polynomials. As $\mathbb{C}[x], \mathcal{S} \subsetneq \mathcal{O}_M$, for the corresponding set of linear functionals we find that $\mathcal{O}'_M \subset (\mathbb{C}[x])' \cap \mathcal{S}'$. Therefore, these distributions give a second appropriate framework. Finally, for a third suitable framework, including the two previous ones, we need to introduce bounded distributions. Let us consider as space of fundamental functions, the linear space $\mathcal{B}$ of bounded smooth functions, i.e., with all its derivatives in $L^\infty(\mathbb{R}^D)$, being the corresponding space of linear functionals $\mathcal{B}'$ the bounded distributions. From $\mathcal{D} \subsetneq \mathcal{B}$ we conclude that bounded distributions are distributions $\mathcal{B}' \subsetneq \mathcal{D}'$. Then, we consider the space of fast decreasing distributions $\mathcal{O}'_c$ given by those distributions $u \in \mathcal{D}'$ such that for each positive integer k, we have $\left(\sqrt{1+x^2}\right)^k u \in \mathcal{B}'$ is a bounded distribution. Any polynomial $P(x) \in \mathbb{C}[x]$, with $\deg P = k$, can be written as

$$P(x) = \left(\sqrt{1+x^2}\right)^k F(x), \qquad F(x) = \frac{P(x)}{\left(\sqrt{1+x^2}\right)^k} \in \mathcal{B}.$$

Therefore, given a fast decreasing distribution $u \in \mathcal{O}'_c$ we may consider

$$\langle u, P(x) \rangle = \left\langle \left(\sqrt{1+x^2}\right)^k u, F(x) \right\rangle$$

which makes sense as $\left(\sqrt{1+x^2}\right)^k u \in \mathcal{B}', F(x) \in \mathcal{B}$. Thus, $\mathcal{O}'_c \subset (\mathbb{C}[x])' \cap \mathcal{D}'$. Moreover it can be proven that $\mathcal{O}'_M \subsetneq \mathcal{O}'_c$, see [113]. Summarizing this discussion, we have found three generalized function spaces suitable for the discussion of polynomials and supports simultaneously:

$$\mathcal{E}' \subset \mathcal{O}'_M \subset \mathcal{O}'_c \subset \left((\mathbb{C}[x])' \cap \mathcal{D}'\right).$$

Definition 1.48 (Distributional sesquilinear forms)**.** *Given a matrix with linear functional as entries*

$$u = \begin{bmatrix} u_{1,1} & \dots & u_{1,p} \\ \vdots & & \vdots \\ u_{p,1} & \dots & u_{p,p} \end{bmatrix}, \qquad u_{i,j} \in (\mathbb{C}[x])',$$

then the matrix of linear functional applies to left and right side of matrix polynomials $P(x)$ are defined as $\langle u, P(x) \rangle$ and $\langle P(x), u \rangle$ respectively, where

$$(\langle u, P(x) \rangle)_{i,j} = \sum_{k=0}^{p} \left\langle u_{i,k}, P_{k,j}(x) \right\rangle \quad \textit{and} \quad (\langle P(x), u \rangle)_{i,j} = \sum_{k=0}^{p} \left\langle u_{k,j}, P_{i,k}(x) \right\rangle.$$

Moreover, the entries of the associated sesquilinear form $\langle P, Q \rangle_u$ are given by

$$\left(\langle P, Q \rangle_u\right)_{i,j} := \sum_{k,l=1}^{p} \left\langle u_{k,l}, P_{i,k}(x) \overline{Q_{j,l}}(x) \right\rangle.$$

When $u_{k,l} \in O_c'$, we write $u \in \left(O_c'\right)^{p\times p}$ and we will say that we have a matrix of fast decreasing distributions. In this case the support is defined as $\operatorname{supp}(u) := \cup_{k,l=1}^{N} \operatorname{supp}(u_{k,l})$.

Remark 1.49. *Notice that the moments of a distributional sesquilinear form are*

$$m_n := \begin{bmatrix} \langle u_{1,1}, x^n \rangle & \dots & \langle u_{1,p}, x^n \rangle \\ \vdots & & \vdots \\ \langle u_{p,1}, x^n \rangle & \dots & \langle u_{p,p}, x^n \rangle \end{bmatrix}$$

and, thus, the block moment matrix has Hankel block structure

$$M := \begin{bmatrix} m_0 & m_1 & m_2 & \cdots \\ m_1 & m_2 & m_3 & \cdots \\ m_2 & m_3 & m_4 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

1.3.1 Gauss–Borel factorization

Following [19]

Proposition 1.50. *If the Gram matrix of a sesquilinear form is quasi-definite, then there exists an unique Gauss–Borel factorization of the moment matrix M given by*

$$M = S_1^{-1} H (S_2)^{-\dagger},$$

where S_1, S_2 are lower unitriangular block matrices and H is a diagonal block matrix. Moreover, if $M = M^\dagger$ then $S_1 = S_2$.

In the sequel, we will use quasi-determinants to obtain connection formulas between some families of orthogonal polynomials. They constitute a generalization of the determinants when the entries of the matrix belong to a non-commutative ring. They share several properties.

Definition 1.51 (Last quasi-determinants [72, 124])**.** *Given a 2×2 block matrix $\left[\begin{smallmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{smallmatrix}\right]$, then*

$$\Theta_* \begin{bmatrix} \boxed{a_{1,1}} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = a_{1,1} - a_{1,2} a_{2,2}^{-1} a_{2,1}, \quad \Theta_* \begin{bmatrix} a_{1,1} & \boxed{a_{1,2}} \\ a_{2,1} & a_{2,2} \end{bmatrix} = a_{1,2} - a_{1,1} a_{2,1}^{-1} a_{2,2},$$
$$\Theta_* \begin{bmatrix} a_{1,1} & \boxed{a_{1,2}} \\ a_{2,1} & a_{2,2} \end{bmatrix} = a_{2,1} - a_{2,2} a_{1,2}^{-1} a_{1,1}, \quad \Theta_* \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & \boxed{a_{2,2}} \end{bmatrix} = a_{2,2} - a_{2,1} a_{1,1}^{-1} a_{1,2},$$

are its quasi-determinants. In particular, $\Theta_* \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & \boxed{a_{2,2}} \end{bmatrix}$ is said to be the last quasi-determinant. When we work with last quasi-determinant, we omit the square in the last entry.

Notice that in each case the quasi-determinant related to the boxed block is just the Schur complement of the opposite block. We will also use quasi-determinants for 3×3 block matrices. In this case, the Sylvester's theorem for quasi-determinants yields

$$\Theta_*\begin{bmatrix}\boxed{a_{1,1}} & a_{1,2} & a_{1,3}\\ a_{2,1} & a_{2,2} & a_{2,3}\\ a_{3,1} & a_{3,2} & a_{3,3}\end{bmatrix}=\Theta_*\begin{bmatrix}\boxed{a_{1,1}} & a_{1,3}\\ a_{3,1} & a_{3,3}\end{bmatrix}-\Theta_*\begin{bmatrix}\boxed{a_{1,2}} & a_{1,3}\\ a_{3,2} & a_{3,3}\end{bmatrix}\Theta_*\begin{bmatrix}\boxed{a_{2,2}} & a_{2,3}\\ a_{3,2} & a_{3,3}\end{bmatrix}^{-1}\Theta_*\begin{bmatrix}\boxed{a_{2,1}} & a_{2,3}\\ a_{3,1} & a_{3,3}\end{bmatrix}, \tag{1.6}$$

when the right side expression makes sense.

Proposition 1.52. *If the last quasi-determinants of the truncated moment matrices are nonsingular, i.e.,*

$$\det\Theta_*(M_{[k]})\neq 0, \qquad\qquad k=1,2,\dots,$$

then the Gauss–Borel factorization exists and the following expressions

$$H_k=\Theta_*\begin{bmatrix}m_0 & \dots & m_{k-1}\\ m_1 & \dots & m_k\\ \vdots & & \vdots\\ m_{k-1} & \dots & m_{2k-2}\end{bmatrix} \qquad (S_1)_{k,l}=\Theta_*\begin{bmatrix}m_0 & m_1 & \dots & m_{k-1} & 0_p\\ \vdots & & \vdots & \vdots & \\ m_{l-1} & m_l & \dots & m_{k+l-2} & 0_p\\ m_l & m_{l+1} & \dots & m_{k+l-1} & I_p\\ m_{l+1} & m_{l+2} & \dots & m_{k+l} & 0_p\\ \vdots & \vdots & & \vdots & \vdots\\ m_k & m_{k+l+1} & \dots & m_{2k-1} & 0_p\end{bmatrix},$$

$$\big((S_2)^\dagger\big)_{k,l}=\Theta_*\begin{bmatrix}m_0 & \dots & m_{l-1} & m_l & m_{l+1} & \dots & m_k\\ m_1 & \dots & m_l & m_{l+1} & m_{l+2} & \dots & m_{k+1}\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ m_{k-1} & \dots & m_{k+l-2} & m_{k+l-1} & m_{k+l} & \dots & m_{2k-1}\\ 0_p & \dots & 0_p & I_p & 0_p & \dots & 0_p\end{bmatrix},$$

hold. We see that the matrices H_k are quasi-determinants. Following [20, 19] we refer to them as quasitau matrices.

1.3.2 Bi-orthogonal polynomials, second kind functions and Christoffel–Darboux kernels

Definition 1.53. *We define*

$$\chi(x):=[I_p,I_px,I_px^2,\dots]^\dagger \quad \textit{and, for}\quad x\neq 0,\quad \chi^*(x):=[I_px^{-1},I_px^{-2},I_px^{-3},\dots]^\dagger.$$

Notice that here we could change $\dagger$ by $\top$ because x is a real variable. It will be used hereinafter.

Remark 1.54. *Observe that the moment matrix can be expressed as*

$$M = \langle \chi(x), \chi(x) \rangle_u . \tag{1.7}$$

Definition 1.55. *Given a quasi-definite matrix of functionals u and the Gauss–Borel factorization* (1.50) *of its Hankel matrix of moments, the corresponding first and second families of matrix bi-orthogonal polynomials are defined by*

$$P^{[1]}(x) = \begin{bmatrix} P_0^{[1]}(x) \\ P_1^{[1]}(x) \\ \vdots \end{bmatrix} := S_1 \chi(x), \qquad P^{[2]}(x) = \begin{bmatrix} P_0^{[2]}(x) \\ P_1^{[2]}(x) \\ \vdots \end{bmatrix} := S_2 \chi(x). \tag{1.8}$$

Definition 1.56. *For $z \notin supp(\mu)$ the corresponding first and second families of second kind functions are defined by*

$$C^{[1]}(z) = \begin{bmatrix} C_0^{[1]}(x) \\ C_1^{[1]}(z) \\ \vdots \end{bmatrix}, \qquad C^{[2]}(z) = \begin{bmatrix} C_0^{[2]}(z) \\ C_1^{[2]}(z) \\ \vdots \end{bmatrix}, \tag{1.9}$$

where for each $n \in \mathbb{N}$,

$$C_n^{[1]}(z) = \left\langle P_n^{[1]}(x), \frac{I_p}{z-x} \right\rangle_u, \qquad C_n^{[2]}(z) = \left\langle \frac{I_p}{z-x}, P_n^{[2]}(x) \right\rangle_u^{\dagger}.$$

Remark 1.57. *The matrix polynomials $P_n^{[i]}(x)$ are monic and $\deg(P_n^{[i]}) = n$, $i = 1,2$.*

Proposition 1.58 (bi-orthogonality). *Given a quasi-definite matrix of linear functionals u, the first and second families of monic matrix polynomials $\left(P_n^{[1]}(x)\right)_{n\in\mathbb{N}}$ and $\left(P_n^{[2]}(x)\right)_{n\in\mathbb{N}}$ are bi-orthogonal*

$$\left\langle P_n^{[1]}(x), P_m^{[2]}(x) \right\rangle_u = \delta_{n,m} H_n, \qquad n,m \in \mathbb{N}. \tag{1.10}$$

Remark 1.59. *These bi-orthogonal relations yield the orthogonality relations*

$$\left\langle P_n^{[1]}(x), x^m I_p \right\rangle_u = 0_p, \qquad \left\langle x^m I_p, P_n^{[2]}(x) \right\rangle_u = 0_p, \qquad m \in \{0, \dots n-1\}, \tag{1.11}$$

$$\left\langle P_n^{[1]}(x), x^n I_p \right\rangle_u = H_n, \qquad \left\langle x^n I_p, P_n^{[2]}(x) \right\rangle_u = H_n. \tag{1.12}$$

Remark 1.60. *If $u = u^{\dagger}$, then $P_n^{[1]}(x) = P_n^{[2]}(x) =: P_n(x)$ and we get an orthogonal set of monic matrix polynomials*

$$\langle P_n(x), P_m(x) \rangle_u = \delta_{n,m} H_n, \qquad n,m \in \mathbb{N}.$$

Some times we will write $\|P_n\|^2 = H_n$. Observe, that in this case, $C_n^{[1]} = C_n^{[2]} =: C_n$.

The shift matrix is the semi-infinite block matrix

$$\Lambda =: \begin{bmatrix} 0_{p\times p} & I_p & 0_{p\times p} & 0_{p\times p} & \cdots \\ 0_{p\times p} & 0_{p\times p} & I_p & 0_{p\times p} & \ddots \\ 0_{p\times p} & 0_{p\times p} & 0_{p\times p} & I_p & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

It satisfies the spectral property

$$\Lambda\chi(x) = x\,\chi(x).$$

If $W(x) = \sum_{k=0}^{N} B_k x^k$, then we define the evaluation of Λ in $W(x)$ as follows

$$W(\Lambda) =: \sum_{k=0}^{N} B_k \Lambda^k.$$

Proposition 1.61. *The symmetry of the block Hankel moment matrix reads as*

$$\Lambda M = M\Lambda^{\dagger}.$$

Notice that this symmetry completely characterizes Hankel block matrices.

Definition 1.62. *The matrices*

$$J_1 := S_1\Lambda S_1^{-1}, \qquad\qquad J_2 := S_2\Lambda S_2^{-1},$$

are said to be the Jacobi matrices associated with the moment matrix M.

Proposition 1.63. *The two Jacobi matrices are related by*

$$H^{-1}J_1 = J_2^{\dagger}H^{-1},$$

being, therefore, block tridiagonal and yield the three recurrence formulas of the bi-orthogonal polynomials and second kind functions

$$\begin{aligned} J_1P^{[1]}(x) &= xP^{[1]}(x), & J_1C^{[1]}(x) &= xC^{[1]}(x) - H_0e_0, \\ J_2P^{[2]}(x) &= xP^{[2]}(x), & J_2C^{[2]}(x) &= xC^{[2]}(x) - H_0^{\dagger}e_0. \end{aligned}$$

where $e_0 = (I_p, 0, 0, \cdots)^{\top}$. When the distributional sesquilinear form satisfies $\mu = \mu^{\dagger}$ we will denote $J_{mon} := J_1 = J_2$ referring to the fact that the orthogonal polynomials are monic.

Proof. The relation between the above two Jacobi matrices follows from the LU factorization and the symmetry $\Lambda M = M\Lambda^{\dagger}$. A consequence of this relation is the three-block-diagonal shape of these matrices. The three term recurrence relations follow from the definitions of the Jacobi matrices in terms of the factorization matrices. ■

Remark 1.64. *Observe that for $i=1,2$, the above relations mean that the sequences $(P_n^{[i]}(x))_{n\in\mathbb{N}}$ satisfy a three term recurrence relation*

$$xP_n^{[i]}(x)=P_{n+1}^{[i]}(x)+a_n^{[i]}P_n^{[i]}(x)+b_n^{[i]}P_{n-1}^{[i]}(x),\quad n\geq 0,\ \ P_0^{[i]}(x)=I_p,\quad P_{-1}^{[i]}(x)=0_{p\times p},$$

where $a_n^{[i]}$, and $b_n^{[i]}$ are matrices of size $p\times p$. Moreover $b_n^{[1]}=H_nH_{n-1}^{-1}$.

Proposition 1.65. *We have the following last quasi-determinantal expressions*

$$P_n^{[1]}(x)=\Theta_*\begin{bmatrix} m_0 & m_1 & \cdots & m_{n-1} & I_p\\ m_1 & m_2 & \cdots & m_n & I_px\\ \vdots & \vdots & & \vdots & \vdots\\ m_{n-1} & m_n & \cdots & m_{2n-2} & I_px^{n-1}\\ m_n & m_{n+1} & \cdots & m_{2n-1} & I_px^n \end{bmatrix},\quad (P_n^{[2]}(x))^\dagger=\Theta_*\begin{bmatrix} m_0 & m_1 & \cdots & m_{n-1} & m_n\\ m_1 & m_2 & \cdots & m_n & m_{n+1}\\ \vdots & \vdots & & \vdots & \vdots\\ m_{n-1} & m_n & \cdots & m_{2n-2} & m_{2n-1}\\ I_p & I_px & \cdots & I_px^{n-1} & I_px^n \end{bmatrix}.$$

Definition 1.66 ([38]). *Given the sequences of matrix monic bi-orthogonal polynomials $(P_n^{[1]}(x),P_n^{[2]}(x))_{n\in\mathbb{N}}$, with respect to the sesquilinear form $\langle\cdot,\cdot\rangle_u$, we define the n-th Christoffel–Darboux kernel matrix polynomial*

$$K_n(x,y):=\sum_{k=0}^{n}(P_k^{[2]}(y))^\dagger H_k^{-1}P_k^{[1]}(x), \tag{1.13}$$

and the mixed Christoffel–Darboux kernel

$$K_n^{(pc)}(x,y):=\sum_{k=0}^{n}\left(P_k^{[2]}(y)\right)^\dagger H_k^{-1}C_k^{[1]}(x).$$

Proposition 1.67. *The Christoffel–Darboux kernel gives the projection operator*

$$\left\langle K_n(x,y),\sum_{j=0}^{\infty}C_jP_j^{[2]}(x)\right\rangle_u=\left(\sum_{j=0}^{n}C_jP_j^{[2]}(y)\right)^\dagger,\quad \left\langle \sum_{j=0}^{\infty}C_jP_j^{[1]}(x),(K_n(y,x))^\dagger\right\rangle_u=\sum_{j=0}^{n}C_jP_j^{[1]}(y), \tag{1.14}$$

where the matrix of functionals u acts on the variable x, while y behaves as a parameter. In particular, we have

$$\left\langle K_n(x,y),I_px^l\right\rangle_u=I_py^l,\quad l\in\{0,1,\dots,n\}. \tag{1.15}$$

Proof. It follows from (1.10). ∎

Proposition 1.68 (Christoffel–Darboux formula)**.** *The Christoffel–Darboux kernel satisfies*

$$(x-y)K_n(x,y)=(P_n^{[2]}(y))^\dagger(H_n)^{-1}P_{n+1}^{[1]}(x)-(P_{n+1}^{[2]}(y))^\dagger(H_n)^{-1}P_n^{[1]}(x),$$

and the mixed Christoffel-Darboux kernel fulfills

$$(x-y)K_n^{(pc)}(x,y)=(P_n^{[2]}(y))^\dagger H_n^{-1}C_{n+1}^{[1]}(x)-(P_{n+1}^{[2]}(y))^\dagger H_n^{-1}C_n^{[1]}(x)+I_p.$$

Proof. We only prove the second formula, since the first one is well known in the literature (see [38]). Indeed, it is a straightforward consequence of the three term recurrence relation. First, let us notice that from Proposition 1.63 and Remark 1.64,

$$\begin{aligned}
xH_k^{-1}C_k^{[1]}(x) - I_p\delta_{k,0} &= H_k^{-1}C_{k+1}^{[1]}(x) + H_k^{-1}a_k^{[1]}C_k^{[1]} + H_k^{-1}b_k^{[1]}C_{k-1}^{[1]}(x) \quad k \geq 0,\\
yP_k^{[2]\dagger}(y)H_k^{-1} &= P_{k+1}^{[2]\dagger}(y)H_{k+1}^{-1}b_{k+1}^{[1]} + P_k^{[2]\dagger}(y)H_k^{-1}a_k^{[1]} + P_{k-1}^{[2]\dagger}(y)H_{k-1}^{-1} \quad k \geq 0.
\end{aligned}$$

From here

$$\begin{aligned}
(x-y)P_k^{[2]\dagger}(y)H_k^{-1}C_k^{[1]}(x) = {} & P_k^{[2]\dagger}(y)H_k^{-1}C_{k+1}^{[1]}(x) + P_k^{[2]\dagger}(y)H_k^{-1}b_k^{[1]}C_{k-1}^{[1]}(x)\\
& - \left[P_{k+1}^{[2]\dagger}(y)H_{k+1}^{-1}b_{k+1}^{[1]}C_k^{[1]}(x) + P_{k-1}^{[2]\dagger}(y)H_{k-1}^{-1}C_k^{[1]}(x)\right] + I_p\delta_{k,0}.
\end{aligned}$$

Summing the later from 0 to n and taking into account that $P_{-1}^{[2]\dagger}(y) = C_{-1}^{[1]}(x) = 0$, the result follows.

■

1.3.3 Positive definite matrix of measures

As in the scalar case, in [96] was proved that a sequence of Hermitian matrices $(m_k)_{k\in\mathbb{N}}$ are the moments of a positive definite matrix of measures $d\mu$ (i.e., the sesquilinear form associated with the measure is non-trivial) if and only if for every sequence of covectors $v_0, v_1 \ldots \in (\mathbb{C}^p)^*$

$$\sum_{i,j}^{k} v_i m_{j+k} v_j^{\dagger} \geq 0,$$

where the equality holds if and only if $v_j = 0,\ j = 0,\ldots,k$.
Given a positive definite matrix of measures $d\mu$, by using a generalization of the Gram-Schmidt orthonormal process for the basis $\{I_p, xI_p, x^2I_p \cdots\}$ of $\mathbb{C}^{p\times p}[x]$, we can to construct sequences of matrix polynomials $(Q_n(x))_{n\in\mathbb{N}}$ with respect to $d\mu$ such that the degree of $Q_n(x)$ is n, their leading coefficients are nonsingular matrices and

$$\int Q_n(x)d\mu Q_k^{\dagger}(x) = I_p\delta_{n,k},$$

where $\delta_{n,k}$ is the Kronecker delta. Observe that the sequence of orthonormal matrix polynomials is not unique since given a sequence of unitary matrices $(U_n)_{n\in\mathbb{N}}$, the sequence $(U_nQ_n(x))_{n\in\mathbb{N}}$ is also a sequence of orthonormal polynomials. If $(P_n(x))_{n\in\mathbb{N}}$ is a sequence of monic polynomials with respect $d\mu$, in particular we can take $Q_n(x) =: \|P_n(x)\|^{-1}P_n(x)$, where $\|P_n(x)\|^2 =: \int P_n(x)d\mu P_n^{\dagger}(x)$. Notice that $\|P_n(x)\|^2$ is a positive definite matrix, thus $\|P_n(x)\|$ is unique. Due to orthogonality of the sequence $(Q_n(x))_{n\in\mathbb{N}}$, it satisfies a three term recurrence relation [38, 134],

$$xQ_n(x) = C_{n+1}Q_{n+1}(x) + E_nQ_n(x) + C_n^{\dagger}Q_{n-1}(x), \quad n \geq 0, \quad Q_{-1}(x) = 0_{p\times p}, \quad Q_0(x) = I_p, \tag{1.16}$$

where C_n are nonsingular matrices and $E_n = E_n^\dagger$, all of them of size $p \times p$. Thus, we can associate with the sequence of matrix polynomials $(Q_n(x))_{n\in\mathbb{N}}$ the following semi-infinite hermitian block matrix

$$J = \begin{bmatrix} E_0 & C_1 & & & \\ C_1^\dagger & E_1 & C_2 & & \\ & C_2^\dagger & E_2 & C_3 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

This matrix is called the Jacobi block matrix associated with the sequences $(Q_n(x))_{n\in\mathbb{N}}$.

Theorem 1.69 (Favard's Theorem [18])**.** *Let $(D_n)_{n\in\mathbb{N}}$ and $(E_n)_{n\in\mathbb{N}}$ be arbitrary matrices with D_n a nonsingular matrix for every $n \geq 1$ and $E_n = E_n^\dagger$. Let $(Q_n(x))_{n\in\mathbb{N}}$ be a sequence of matrix polynomials defined by the recurrence formula*

$$xQ_n(x) = D_{n+1}Q_{n+1}(x) + E_nQ_n(x) + D_n^\dagger Q_{n-1}(x), \qquad n \geq 0, \quad Q_0(x) = I_p, \quad Q_{-1}(x) = 0_{p\times p}.$$

Then, there exists a Hermitian matrix of measures $d\mu$ such that $(Q_n(x))_{n\in\mathbb{N}}$ is the sequence of matrix orthonormal polynomials with respect to $d\mu$.

Chapter 2

Generalization of the Favard's theorem, higher order recurrence relations and connection with matrix orthogonal polynomials

In the scalar case, the Favard's theorem is one of the most important results in the theory of orthogonal polynomials since given sequences of real numbers $(c_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$, with $c_n \neq 0$ for every $n \in \mathbb{N}$, and the sequence of polynomials $(p_n(x))_{n\in\mathbb{N}}$ defined by

$$xp_n(x) = c_{n+1}p_{n+1}(x) + \lambda_n p_n(x) + c_n p_{n-1}(x), \quad p_{-1}(x) = 0, \quad p_0(x) = 1,$$

there exists an unique symmetric bilinear form $B(\cdot,\cdot)$ (Definition 1.45) such that $B(p_n(x), p_m(x)) = \delta_{n,m}$ and $B(xf,g) = B(f,xg)$ for every polynomial f,g. Moreover, there exists a function μ given from a function $g(x)$ in the Schwartz space (see [51, 50]) (so a $\mathcal{C}^\infty$-function) as follows

$$\mu(x) = \int_{-\infty}^{x} g(t)dt \tag{2.1}$$

such that $B(x^n,x^m) = \int x^{n+m}d\mu$. Note that the converse is also true. However, in many contributions, for particular symmetric bilinear forms such that the property $B(xf,g) = B(f,xg)$ is not satisfied, the authors have found that the corresponding sequences of orthogonal polynomials satisfy higher order recurrence relations (see for example [8, 65, 84, 90]). Thus, it is natural to ask for the generalization of the Favard theorem, i.e. given a sequence of scalar polynomials satisfying a higher order recurrence relation

$$x^N p_n(x) = c_{n,0}p_n(x) + \sum_{k=1}^{N} \left[c_{n,k}p_{n-k}(x) + c_{n+k,k}p_{n+k}(x)\right],$$

where N is a fixed nonnegative integer and $(c_{n,N})_{n\in\mathbb{N}}$ is a sequence of nonzero real numbers and $(c_{n,k})_{n\in\mathbb{N}}$, with $1 \leq k \leq N$, are sequences of real numbers, there exists a symmetric bilinear form

$B(\cdot,\cdot)$ such that $(p_n(x))_{n\in\mathbb{N}}$, is a sequence of orthonormal polynomials with respect to $B(\cdot,\cdot)$ and if it exists , how can be represented?. The converse problem is also satisfied?.
In this first section the main results of A. J. Durán in [51] about polynomials satisfying such a kind of higher order recurrence relations are summarized in order to provide the answer to the above questions.

2.1 Generalization of the Favard's theorem

Definition 2.1 ([51]). *Let $B(\cdot,\cdot)$ be a symmetric bilinear form such that the multiplication operator by x^N is symmetric for $B(\cdot,\cdot)$, and let w be a primitive N-th root of the unity*[1]. *Given $0 \leq m \leq N-1$, we define the operator $T_{m,N} : \mathbb{R}[x] \to \mathbb{R}[x]$ as follows*

$$T_{m,N}(f)(x) = \frac{1}{N}\sum_{k=0}^{N-1}(w^{-m})^k f(w^k x).$$

Remark 2.2. *For a polynomial $f(x) = \sum_i a_i x^i$, from the properties of a primitive root of the unity we get*

$$T_{m,N}(f)(x) = \sum_i a_{iN+m}x^{iN+m}.$$

Theorem 2.3 ([51]). *Let $B(\cdot,\cdot)$ be a symmetric bilinear form. Then the following statements are equivalent.*

i) *The multiplication operator by x^N is symmetric with respect to $B(\cdot,\cdot)$, that is, $B(x^N f,g) = B(f,x^N g)$ for every polynomial f,g.*

ii) *There exist functions μ_0 and $\mu_{m,m'}$ as in (2.1) with $\mu_{m,m'} = \mu_{m',m}$, for $1 \leq m,m' \leq N-1$, such that $B(\cdot,\cdot)$ can be written as follows*

$$B(f,g) = \int f(x)g(x)d\mu_0 + \sum_{1\leq m,m'\leq N-1}\int T_{m.N}(f)(x)T_{m',N}(g)(x)d\mu_{m,m'}.$$

iii) *There exist functions $\mu_{m,m'}$ as in (2.1), with $\mu_{m,m'} = \mu_{m',m}$, for $0 \leq m,m' \leq N-1$, such that $B(\cdot,\cdot)$ can be written as follows*

$$B(f,g) = \sum_{0\leq m,m'\leq N-1}\int T_{m.N}(f)(x)T_{m',N}(g)(x)d\mu_{m,m'}.$$

With this result it is possible to extend the Favard's theorem when a sequence of the polynomials $(p_n(x))_{n\in\mathbb{N}}$ satisfies a $(2N+1)$-term recurrence relation formula

[1] We said that w be a primitive N-th roots of the unity if $w^N = 1$ for N a positive integer and for all integer n such that $0 < n < N$ we have that $w^n \neq 1$

$$x^N p_n(x) = c_{n,0} p_n(x) + \sum_{k=1}^{N} \left[c_{n,k} p_{n-k}(x) + c_{n+k,k} p_{n+k}(x) \right], \tag{2.2}$$

with the convention $p_k(x) = 0$ for $k < 0$. Here $(c_{n,N})_{n \in \mathbb{N}}$ is a sequence of nonzero real numbers and $(c_{n,k})_{n \in \mathbb{N}}$, with $1 \leq k \leq N$, are sequences of real numbers. Notice that for $N = 1$ the usual three term recurrence relation given in (1.16) follows. Now we can state the following theorem.

Theorem 2.4 ([51]). *Let N be a positive integer and $(p_n(x))_{n \in \mathbb{N}}$ be a sequence of polynomials satisfying the $(2N+1)$-term recurrence relation (2.2). Then, there exist functions μ_0 and $\mu_{m,m'}$ as in (2.1), $1 \leq m, m' \leq N-1$, with $\mu_{m,m'} = \mu_{m',m}$, such that the sequence of polynomials $(p_n(x))_{n \in \mathbb{N}}$ is orthogonal with respect to a symmetric quasi-definite bilinear form $B(\cdot,\cdot)$ defined as follows*

$$B(f,g) = \int f(x) g(x) d\mu_0 + \sum_{1 \leq m, m' \leq N-1} \int T_{m.N}(f)(x) T_{m',N}(g)(x) d\mu_{m,m'}.$$

Let us suppose now that the multiplication operator by a polynomial $h(x)$ of degree N is symmetric with respect to a symmetric bilinear form $B(\cdot,\cdot)$, i.e. $B(hf,g) = B(f,gh)$. The idea is to extend Theorems 2.3 and 2.4 changing x^N by the polynomial $h(x)$. For this we are going to take a new basis in the polynomial space $\mathbb{R}[x]$ as follows

$$\mathfrak{B}_h =: \{x^m h^k(x),\ k \geq 0, 0 \leq m \leq N-1\}.$$

Note that every polynomial f can be written as $f(x) = \sum_{k \geq 0} \sum_{m=0}^{N-1} a_{m,k} x^m h^k(x)$, where $a_{m,k}$ are the coefficients with respect to the new basis. Now, we define the operator $T_{m,h}(f)(x)$ as follows

$$T_{m,h}(f)(x) = \sum_{k \geq 0} a_{m,k} x^m h^k(x).$$

Taking into account the above definition we can state the following

Theorem 2.5 ([51]). *Let $B(\cdot,\cdot)$ be a real symmetric bilinear form. The following statements are equivalent.*

i) *The multiplication operator by $h(x)$ is symmetric with respect to $B(\cdot,\cdot)$, that is, $B(hf,g) = B(f,hg)$ for every polynomial f,g.*

ii) *There exist functions μ_0 and $\mu_{m,m'}$ as in (2.1), $1 \leq m, m' \leq N-1$, with $\mu_{m,m'} = \mu_{m',m}$, such that $B(\cdot,\cdot)$ can be written as follows*

$$B(f,g) = \int f(x) g(x) d\mu_0 + \sum_{1 \leq m \leq m' \leq N-1} \int T_{m.h}(f)(x) T_{m',h}(g)(x) d\mu_{m,m'}.$$

Notice that if in the recurrence relation (2.2) we change the term x^N by $h(x)$, we can extend Theorem 2.4 from the above theorem. So, A. J. Durán in [51] has generalized the Favard's theorem not only for monomials of the form x^N with $N \in \mathbb{N}$, but for any polynomial $h(x)$.

2.1.1 Inner products of Sobolev type

Recall that the diagonal inner products of Sobolev type are defined as

$$B(f,g) = \int f(x)g(x)d\mu + \sum_{l=0}^{N-1} \int f^{(l)}(x)g^{(l)}(x)d\mu_l. \tag{2.3}$$

One of the most interesting problems is to study the inner products of Sobolev type when the measures μ_l are Dirac's deltas supported at the same point (see [8, 65, 84, 90]), i.e. when

$$B(f,g) = \int f(x)g(x)d\mu + \sum_{l=0}^{N-1} M_l f^{(l)}(a)g^{(l)}(a).$$

Given a real symmetric bilinear form $B(\cdot,\cdot)$, not necessarily as (2.3), we want to find necessary and sufficient conditions for $B(\cdot,\cdot)$ such that it can be written as

$$B(f,g) = \int f(x)g(x)d\mu + \sum_{l,l'=1}^{K} \sum_{i=0}^{n_l-1} \sum_{j=0}^{n_{l'}-1} M_{i,j,l',l} f^{(i)}(a_l)g^{(j)}(a_{l'}),$$

where μ is a function as in (2.1) and $M_{i,j,l',l}$ are positive constants. For special cases, we can find conditions for $B(\cdot,\cdot)$ such that it can be represented as in (2.3), when the measures μ_l are Dirac's deltas at the same point, or when $\mu_l = 0$, for $l \neq 1$, and μ_1 is a finite combination of Dirac deltas.

Theorem 2.6 ([51]). *Let $B(\cdot,\cdot)$ be a real symmetric bilinear form defined in $\mathbb{R}[x]$ and N a positive integer. The following statements are equivalent.*

i) *The multiplication operator by x^N is symmetric with respect $B(\cdot,\cdot)$ and $B(x^N f, xg) = B(xf, x^N g)$ for all polynomials f,g.*

ii) *There exist a function μ as in (2.1) and constants $M_{k,m}$, $1 \leq m,k \leq N-1$, $M_{m,k} = M_{k,m}$, such that $B(\cdot,\cdot)$ can be written as follows*

$$B(f,g) = \int f(x)g(x)d\mu + \sum_{k,m=1}^{N-1} M_{k,m} f^{(k)}(0)g^{(m)}(0).$$

Moreover, if in condition (i), $B(\cdot,\cdot)$ satisfies $B(x^k,x^m) = B(1,x^{k+m},)$ when $1 \leq k,m \leq N-1$ and $k \neq m$, this is equivalent to the fact that there exist a function μ as in (2.1) and constants M_k, $1 \leq k \leq N-1$, such that $B(\cdot,\cdot)$ can be written as follows

$$B(f,g) = \int f(x)g(x)d\mu + \sum_{k=1}^{N-1} M_k f^{(k)}(0)g^{(k)}(0).$$

Notice that the above theorem characterizes the symmetric bilinear form (2.3) when μ_l are Dirac deltas supported at zero.

The following two lemmas show other characterization of the symmetric bilinear forms when some specific properties with respect to a polynomial $h(x)$ are satisfied by the bilinear form.

Lemma 2.7 ([51]). *Let $B(\cdot,\cdot)$ be a real symmetric bilinear form defined in $\mathbb{R}[x]$ and K a non-negative integer number. Let us consider a finite sequence of real numbers $(a_l)_{l=1}^K$ and non-negative integers n_l, $1 \le l \le K$. Let $h(x)$ be the polynomial $h(x) = (x-a_1)^{n_1}\cdots(x-a_K)^{n_K}$ and $N = n_1 + \cdots n_K = deg(h)$. Then the following statements are equivalent.*

i) *If f,g are polynomials, then $B(hf,g) = 0$.*

ii) *There exist constants $M_{i,j,l',l}$, $0 \le i \le n_l - 1$, $0 \le j \le n_{l'} - 1$, $1 \le l,l' \le K$, and $M_{i,j,l',l} = M_{j,i,l,l'}$, such that $B(\cdot,\cdot)$ is defined by*

$$B(f,g) = \sum_{l,l'=1}^{K} \sum_{i=0}^{n_l-1} \sum_{j=0}^{n_{l'}-1} M_{i,j,l',l} f^{(i)}(a_l) g^{(j)}(a_{l'}).$$

Theorem 2.8 ([51]). *Let $B(\cdot,\cdot)$ be a real symmetric bilinear form in $\mathbb{R}[x]$ and K a non-negative integer. Let us consider a finite sequence of real numbers $(a_l)_{l=1}^K$ and non-negative integers n_l, $1 \le l \le K$. Let $h(x)$ be the polynomial $h(x) = (x-a_1)^{n_1}\cdots(x-a_K)^{n_K}$ and $N = n_1 + \cdots n_K = deg(h)$. Then the following statements are equivalent.*

i) *The multiplication operator by $h(x)$ is symmetric with respect to $B(\cdot,\cdot)$ and $B(hf,xg) = B(xf,hg)$.*

ii) *There exist a function μ as in (2.1) and constants $M_{i,j,l',l}$ with $0 \le i \le n_l - 1$, $0 \le j \le n_{l'} - 1$, $1 \le l,l' \le K$, and $M_{i,j,l',l} = M_{j,i,l,l'}$, such that*

$$B(f,g) = \int f(x)g(x)d\mu + \sum_{l,l'=1}^{K} \sum_{i=0}^{n_l-1} \sum_{j=0}^{n_{l'}-1} M_{i,j,l',l} f^{(i)}(a_l) g^{(j)}(a_{l'}).$$

2.2 Connection between sequences of scalar orthonormal polynomials and matrix orthonormal polynomials

Doing small changes in the proof of Theorem 2.4 it is possible show that if $(p_n(x))_{n\in\mathbb{N}}$ is a sequence of polynomials satisfying a $(2N+1)$-recurrence formula

$$h(x)p_n(x) = c_{n,0}p_n(x) + \sum_{k=1}^{N} [\overline{c}_{n,k}p_{n-k}(x) + c_{n+k,k}p_{n+k}(x)],$$

with the convention $p_k(x) = 0$, for $k < 0$, and $h(x)$ is a monic polynomial of degree N, $(c_{n,0})_{n\in\mathbb{N}}$ is a real sequence, $(c_{n,k})_{n\in\mathbb{N}}$ are complex sequences for $1 \le k \le N$, with $c_{n,N} \ne 0$ for every $n \in \mathbb{N}$, then there exists a positive definite matrix of measures $M = (\mu_{k,l})_{k,l=0}^{N-1}$ (see [54]) such that $(p_n(x))_{n\in\mathbb{N}}$ is orthogonal with respect to the inner product

$$B_\mu(p,q) = \sum_{k,j=0}^{N-1} \int R_{h,k}(p)\overline{R_{h,j}(q)}d\mu_{k,j}, \tag{2.4}$$

where the operator $R_{h,k}(p)(x)$ is defined as follows

Definition 2.9. *Let $h(x)$ be a monic polynomial of degree N and $g(x)$ be a polynomial given in terms of the basis $\{x^k h(x)^m; k = 0,1,\dots,N-1, m \in \mathbb{N}\}$ i.e.*

$$g(x) = \sum_{m=0} \sum_{k=0}^{N-1} b_{k,n} x^k h^m(x).$$

Then we define the linear operator $R_{h,k}(g)$ as

$$R_{h,k}(g)(x) = \sum_{m=0} b_{k,n} x^m.$$

It should be noted that the operator $R_{h,k}(g)(x)$ takes from g just those powers with remainder k, modulus N, and then removes x^k and changes $h(x)$ to x. Thus

$$g(x) = R_{h,0}(g)(h(x)) + xR_{h,1}(g)(h(x)) + \cdots + x^{N-1} R_{h,N-1}(g)(h(x)).$$

We can formulate the following question: what is the connection between the sequence of the matrix orthonormal polynomials with respect to M and the sequence of polynomials $(p_n(x))_{n\in\mathbb{N}}$?. The main theorem in this subsection, which appears in [64], gives an answer to the question. There the following result is proved. Given a sequence of scalar orthonormal polynomials $(p_n(x))_{n\in\mathbb{N}}$ satisfying a higher order recurrence relation it is always possible to find a sequence of matrix orthonormal polynomials related to $(p_n(x))_{n\in\mathbb{N}}$ such that it satisfies a three term recurrence relation like (1.16).

Theorem 2.10 ([64])**.** *Let us assume that $(p_n(x))_{n\in\mathbb{N}}$ is a sequence of scalar orthonormal polynomials satisfying the following $(2N+1)$-term recurrence relation*

$$h(x)p_n(x) = c_{n,N} p_n(x) + \sum_{k=1}^{N} [\bar{c}_{n,k} p_{n-k}(x) + c_{n+k,k} p_{n+k}(x)], \quad \textit{with} \quad p_k(x) = 0 \ \textit{for} \ \ k < 0, \tag{2.5}$$

where $(c_{n,0})_{n\in\mathbb{N}}$ is sequence of real numbers and $(c_{n,k})_{n\in\mathbb{N}}$ are sequences of complex numbers for $1 \le k \le N$ with $c_{n,N} \neq 0$ for all $n \in \mathbb{N}$. Let us define the sequence of matrix polynomials $(P_n(x))_{n\in\mathbb{N}}$ by

$$P_n(x) = \begin{bmatrix} R_{h,0}(p_{nN})(x) & R_{h,1}(p_{nN})(x) & \cdots & R_{h,N-1}(p_{nN})(x) \\ R_{h,0}(p_{nN+1})(x) & R_{h,1}(p_{nN+1})(x) & \cdots & R_{h,N-1}(p_{nN+1})(x) \\ \vdots & \vdots & & \vdots \\ R_{h,0}(p_{nN+N-1})(x) & R_{h,1}(p_{nN+N-1})(x) & \cdots & R_{h,N-1}(p_{nN+N-1})(x) \end{bmatrix}.$$

Then the sequence of matrix polynomials $(P_n(x))_{n\in\mathbb{N}}$ is orthonormal on the real line with respect to a positive definite matrix of measures and it satisfies a three term recurrence relation with matrix coefficients

$$xP_n(x) = D_{n+1}P_{n+1}(x) + E_nP_n(x) + D_n^{\dagger}P_{n-1}(x), \tag{2.6}$$

where D_n is a nonsingular lower triangular matrix and $E_n = E_n^{\dagger}$. Moreover

$$E_n = \begin{bmatrix} c_{nN,0} & c_{nN+1,1} & \cdots & c_{nN+N-1,N-1} \\ \overline{c}_{nN+1,1} & c_{nN+1,0} & \cdots & c_{nN+N-1,N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{c}_{nN+N-1,N-1} & \overline{c}_{nN+N-1,N-2} & \cdots & c_{nN+N-1,0} \end{bmatrix}, \quad D_n = \begin{bmatrix} c_{nN,N} & & & \\ c_{nN,N-1} & c_{nN+1,N} & & \\ \vdots & \vdots & \ddots & \\ c_{nN,1} & c_{nN+1,2} & \cdots & c_{nN+N-1,N} \end{bmatrix}. \tag{2.7}$$

Conversely let us assume that $(P_n(x))_{n\in\mathbb{N}}$ is a sequence of orthonormal matrix polynomials or, equivalently, they satisfy a three term recurrence relation like (2.6), where without loss of generality we can suppose that $P_n(x)$ has as leading coefficient a lower triangular matrix. If we denote by $(P_n)_{j,m}(x)$ the (j,m) entry of $P_n(x)$, then the sequence of polynomials $(p_n(x))_{n\in\mathbb{N}}$ defined by

$$p_{nN+j}(x) = \sum_{m=0}^{N-1} x^m (P_n)_{j,m}(h(x)) \tag{2.8}$$

satisfies a $(2N+1)$ term recurrence relation.

Proof. First, we will see that if $(p_n(x))_{n\in\mathbb{N}}$ satisfies (2.5), then $(P_n(x))_{n\in\mathbb{N}}$ is a sequence of matrix orthonormal polynomials. The proof of the converse result is similar. Since $(p_n(x))_{n\in\mathbb{N}}$ is a sequence of orthonormal polynomials associated with a Hermitian bilinear form (2.4), we can take the following matrix of measures

$$dM = \begin{bmatrix} d\mu_{0,0} & d\mu_{0,1} & \cdots & d\mu_{0,N-1} \\ \vdots & \vdots & & \vdots \\ d\mu_{N-1,0} & d\mu_{N-1,1} & \cdots & d\mu_{N-1,N-1} \end{bmatrix},$$

and, for $0 \leq i,m, \leq N-1$,

$$\int \begin{bmatrix} R_{h,0}(p_{nN+i}) & \cdots & R_{h,N-1}(p_{nN+i}) \end{bmatrix} dM \begin{bmatrix} \overline{R_{h,0}}(p_{nN+m}) \\ \vdots \\ \overline{R_{h,N-1}}(p_{nN+m}) \end{bmatrix}$$
$$= \sum_{k,j=0}^{N-1} \int R_{h,k}(p_{nN+i})\overline{R_{h,j}}(p_{nN+m})d\mu_{k,j} = \delta_{i,m}.$$

Then, the above expression yields

$$\int P_n(x)dMP_m^{\dagger}(x) = \delta_{n,m}I_N.$$

Let us see now that $(P_n(x))_{n\in\mathbb{N}}$ satisfies a three term recurrence relation by giving in an explicit way the matrices D_n and E_n in (2.6). Since $(p_n(x))_{n\in\mathbb{N}}$ satisfies (2.5), then we can associate with it a $(2N+1)$-banded infinite Hermitian matrix J,

$$J=\begin{bmatrix} c_{0,0} & c_{1,1} & \cdots & c_{N-1,N-1} & c_{N,N} & & & & & \\ \overline{c}_{1,1} & c_{1,0} & \cdots & \cdots & c_{N,N-1} & c_{N+1,N} & & & & \\ \vdots & \vdots & \ddots & & & \ddots & \ddots & & & \\ \overline{c}_{N,N} & \overline{c}_{N,N-1} & \cdots & \ddots & & & c_{2N-1,N-1} & c_{2N,N} & & \\ 0 & \overline{c}_{N+1,N} & \cdots & & \ddots & & & c_{2N,N-1} & c_{2N+1,N} & \\ & & \ddots & & & & & & \ddots & \ddots \end{bmatrix},$$

or, equivalently,

$$J=\begin{bmatrix} E_0 & D_1 & & & \\ D_1^{\dagger} & E_1 & D_2 & & \\ & D_2^{\dagger} & E_2 & D_3 & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

where the matrices E_i and D_i, ($i\in\mathbb{N}$) are defined in (2.7). Now, if we compute the (k,m) entry of the matrix

$$D_{n+1}P_{n+1}(x)+E_nP_n(x)+D_n^{\dagger}P_{n-1}(x),$$

then we have

$$\begin{aligned}
&\sum_{j=0}^{k} c_{(n+1)N+j,N+j-k}R_{h,m}\left(p_{(n+1)N+j}\right)(x)+\sum_{j=k}^{N-1} c_{Nn+j,|k-j|}R_{h,m}\left(p_{nN+j}\right)(x)\\
&\qquad+\sum_{j=0}^{k-1}\overline{c}_{Nn+k,|k-j|}R_{h,m}\left(p_{nN+j}\right)(x)+\sum_{j=k}^{N-1}\overline{c}_{Nn+k,N+k-j}R_{h,m}\left(p_{(n-1)N+j}\right)(x)\\
&\qquad=\sum_{l=N-k}^{N} c_{Nn+l+k,l}R_{h,m}\left(p_{nN+l+k}\right)(x)+\sum_{l=0}^{N-k-1} c_{nN+k+l,l}R_{h,m}\left(p_{nN+l+k}\right)(x)\\
&\qquad\quad+\sum_{l=k}^{1}\overline{c}_{nN+k,l}R_{h,m}\left(p_{nN+k-l}\right)(x)+\sum_{l=N}^{k+1}\overline{c}_{Nn+k,l}R_{h,m}\left(p_{nN+k-l}\right)(x)\\
&=c_{nN,0}R_{h,m}(p_{nN+k})(x)+\sum_{l=1}^{N}[\overline{c}_{nN+k,l}R_{h,m}(p_{nN+k-l})(x)+c_{nN+k+l,l}R_{h,m}(p_{nN+k+l})(x)].
\end{aligned}\tag{2.9}$$

According to the linearity of the operator $R_{h,m}(\cdot)$ and (2.5), for $0\le k,m\le N-1$ we have

$$xR_{h,m}(p_{Nn+k})(x)=c_{nN+k,0}R_{h,m}\left(p_{nN+k}\right)(x)+\sum_{l=1}^{N}[c_{nN+k+l,l}R_{h,m}\left(p_{nN+k+l}\right)(x)+\overline{c}_{nN+k,l}R_{h,m}\left(p_{nN+k-l}\right)(x)]$$

and this implies (2.6).

As D_n is a nonsingular lower triangular matrix, then D_n^{-1} is also a lower triangular matrix. Using this fact and that $P_{-1}(x) = 0_{N\times N}$, $P_0(x) = I_N$, we can conclude by a recurrence argument that $P_n(x)$ has always a lower triangular matrix as leading coefficient.

Conversely, suppose that $(P_n(x))_{n\in\mathbb{N}}$, where $P_n(x) = (P_n)_{j,m=0}^{N-1}$, is a sequence of orthonormal matrix polynomials or, equivalently, satisfying a three term recurrence relation as in (2.6). Notice that in order to the polynomials defined in (2.8) have degree $nN + j$ we need that $(P_n)_{j,m}(h(x)) = 0$ for $m > j$, and we can not guarantee this up to $P_n(x)$ has as leading coefficient a lower triangular matrix. We will see that it is always possible by assuming that both D_n in (2.6) and the leading coefficient of $P_n(x)$ are lower triangular matrices.

The sequence of matrix orthonormal polynomials $(P_n(x))_{n\in\mathbb{N}}$ with respect to a matrix of measures M is not unique in the sense, that if $(U_n)_{n\in\mathbb{N}}$ is a family of unitary matrices, then $(U_nP_n(x))_{n\in\mathbb{N}}$ is also a sequence of orthonormal polynomials with respect to the matrix of measures M. Suppose that $(S_n(x))_{n\in\mathbb{N}}$ is a sequence of matrix orthonormal polynomials satisfying the following three term recurrence relation,

$$\begin{aligned} &xS_n(x) = B_{n+1}S_{n+1}(x) + A_nS_n(x) + B_nS_{n-1}(x), \qquad n \geq 0,\\ &S_{-1}(x) = 0_{N\times N}, \qquad S_0(x) = I_N, \end{aligned}$$

then the polynomials $P_n(x) = U_nS_n(x)$ satisfy

$$\begin{aligned} &xP_n(x) = U_nB_{n+1}U_{n+1}^{\dagger}P_{n+1}(x) + U_nA_nU_n^{\dagger}P_n(x) + U_nB_n^{\dagger}U_{n-1}^{\dagger}P_{n-1}(x), \quad n \geq 0,\\ &P_{-1}(x) = 0_{N\times N}, \qquad P_0(x) = I_N. \end{aligned}$$

Can U_n be chosen such that $L_n = U_{n-1}B_nU_n^{\dagger}$ is a lower triangular matrix?. The answer is yes. For see this, we are going to use the following argument.
We choose an unitary matrix U_0 such that $U_0S_0 = P_0$ is a lower triangular matrix. Next, let us notice that the matrix U_0B_1 can be factorized as[2] $U_0B_1 = D_1U_1$, where D_1 is a lower triangular matrix and U_1 is an unitary matrix. Thus $D_1 = U_0S_1U_1^{\dagger}$. If $U_0, U_1, \cdots U_{n-1}$, are given, then we can find U_n such that $U_{n-1}B_n = D_nU_n$, with D_n a lower triangular matrix. Thus we can always assume that both the leading coefficient of $P_n(x)$ and the sequence D_n can be chosen as lower triangular matrices. If $(P_n(x))_{n\in\mathbb{N}}$ satisfies (2.6), where D_n and E_n are as in (2.7), then the entry (k,m) of $P_n(x)$ satisfies

[2]**QR** factorization of J. Francis and N. Kublanovskaya: Let A be a nonsingular matrix then is can written as

$$A = RQ$$

where R is a lower triangular matrix and U an unitary matrix

$$
\begin{aligned}
h(P_n)_{k,m}(h) = & \sum_{j=0}^{k} c_{(n+1)N+j,N+j-k}(P_{n+1})_{j,m}(h) + \sum_{j=k}^{N-1} c_{Nn+j,|k-j|}(P_n)_{j,m}(h) \\
& + \sum_{j=0}^{k-1} \overline{c}_{Nn+k,|k-j|}(P_n)_{j,m}(h) + \sum_{j=k}^{N-1} \overline{c}_{Nn+k,N+k-j}(P_{n-1})_{j,m}(h).
\end{aligned}
$$

We can multiply both sides of the above expression by x^m and as the index in the summation (2.8) does not depend on n, we have

$$
\begin{aligned}
h(x)p_{nN+j}(x) = & \sum_{j=0}^{k} c_{(n+1)N+j,N+j-k}p_{(n+1)N+j}(x) + \sum_{j=k}^{N-1} c_{Nn+j,|k-j|}p_{nN+j}(x) \\
& + \sum_{j=0}^{k-1} \overline{c}_{Nn+k,|k-j|}p_{nN+j}(x) + \sum_{j=k}^{N-1} \overline{c}_{Nn+k,N+k-j}p_{(n-1)N+j}(x).
\end{aligned}
$$

A similar argument as in (2.9) shows that $(p_n(x))_{n\in\mathbb{N}}$ satisfies a $(2N+1)$-term recurrence relation. ■

2.2.1 Example

2.11 Example ([64]). *Let us consider the following discrete Sobolev inner product*

$$B(f,g) = \int fgd\mu + \sum_{i=0}^{M}\sum_{j=0}^{M_i} C_{i,j} f^{(j)}(x_i)g^{(j)}(x_i), \tag{2.10}$$

where f,g are polynomials and $C_{i,j}$ are nonnegative real numbers. Take $N = M + \sum_{i=1}^{M} M_i$. Let $h(x)$ be the polynomial

$$h(x) = \prod_{i=0}^{M}(x-x_i)^{M_i+1}.$$

It is clear that $h(x)$ satisfies $B(hf,g) = B(f,hg)$. Then from Theorem 2.4 changing $x^N$for $h(x)$ (see[51]), the sequence of orthonormal polynomials with respect to $B(\cdot,\cdot)$ satisfies a $2N+1$-term recurrence relation as in (2.5). Given a polynomial g of degree $sN+l$ and the basis $\{x^i h^k : \ 0 \le i \le N-1, \ \ k \ge 0\}$, let us write

$$g(x) = \sum_{k=0}^{s}\sum_{n=0}^{N-1} a_{n,k}x^n h^k,$$

where $a_{m,s} = 0$ for $m > l$. According to the representation given in Definition 2.9, we get

$$g(x) = \sum_{n=0}^{N-1} x^n R_{h,n}(g)(h(x)) = \begin{bmatrix} R_{h,0}(g)(h) & \cdots & R_{h,N-1}(g)(h) \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ x^{N-1} \end{bmatrix}.$$

In this sense, we can say that the polynomial g is equivalent to the polynomial vector $[R_{h,0}(g)(h), \cdots, R_{h,N-1}(g)(h)]$. Using the above representation for $g(x)$ we can compute $g^{(j)}(x)$ as follows

$$g^{(j)}(x) = \sum_{n=j}^{N-1} \sum_{q=0}^{j} \binom{j}{q} \frac{n!}{(n-q)!} x^{n-q} \frac{\partial^{j-q}}{\partial x^{j-q}} R_{h,n}(g)(h),$$

and from the Faá di Bruno formula (see [30])

$$\frac{\partial^m}{\partial x^m} R_{h,n}(g)(h(x)) = \sum_{i=0}^{m} \frac{\partial^i}{\partial h^i} R_{h,k}(g)(h(x)) \times \sum_{a_1,\cdots a_m} \binom{m}{a_1,\cdots a_m} \times \left(\frac{h^{(1)}(x)}{1!}\right)^{a_1} \cdots \left(\frac{h^{(m)}(x)}{m!}\right)^{a_m},$$

where $a_1 + a_2 + \cdots + a_m = i$ and $a_1 + 2a_2 + \cdots + ma_m = m$. From here, for $i = 0, \cdots, m$,

$$\frac{\partial^m}{\partial x^m} R_{h,n}(g)(h(x_i)) = 0, \qquad 1 \le m \le M_i.$$

Notice that $g^{(j)}(x_i)$ can be written as

$$g^{(j)}(x_i) = \sum_{n=j}^{N-1} \frac{n!}{(n-j)!} x_i^{n-j} R_{h,n}(g)(0).$$

From the above analysis, the inner product (2.10) can be written as

$$B(f,g) = \int \begin{bmatrix} R_{h,0}(f)(h) & \cdots & R_{h,N-1}(f)(h) \end{bmatrix} dM(x) \begin{bmatrix} R_{h,0}(g)(h) \\ \vdots \\ R_{h,N-1}(g)(h) \end{bmatrix}$$
$$+ \begin{bmatrix} R_{h,0}(f)(0) & \cdots & R_{h,N-1}(f)(0) \end{bmatrix} L \begin{bmatrix} R_{h,0}(g)(0) \\ \vdots \\ R_{h,N-1}(g)(0) \end{bmatrix},$$

where

$$dM(x) = \begin{bmatrix} d\mu & \cdots & x^{N-1}d\mu \\ \vdots & & \vdots \\ x^{N-1}d\mu & \cdots & x^{2N-2}d\mu \end{bmatrix},$$

and $L = \sum_{i=0}^{M} \sum_{j=0}^{M_i} C_{i,j} L(i,j)$ *with*

$$L(i,j) = \begin{bmatrix} 0 \\ \vdots \\ j! \\ \vdots \\ \frac{(N-1)!}{(N-1-j)!} x_i^{N-1-j} \end{bmatrix} \times \begin{bmatrix} 0 & \cdots & j! & \cdots & \frac{(N-1)!}{(N-1-j)!} x_i^{N-1-j} \end{bmatrix}.$$

Thus, from Theorem 2.10, if $(p_n(x))_{n\in\mathbb{N}}$ *is the sequence of orthonormal polynomials with respect to the inner product (2.10), then the sequence of matrix polynomials* $(P_n(x))_{n\in\mathbb{N}}$ *defined by*

$$P_n(x) = \begin{bmatrix} R_{h,0}(p_{nN})(x) & R_{h,1}(p_{nN})(x) & \cdots & R_{h,N-1}(p_{nN})(x) \\ R_{h,0}(p_{nN+1})(x) & R_{h,1}(p_{nN+1})(x) & \cdots & R_{h,N-1}(p_{nN+1})(x) \\ \vdots & \vdots & & \vdots \\ R_{h,0}(p_{nN+N-1})(x) & R_{h,1}(p_{nN+N-1})(x) & \cdots & R_{h,N-1}(p_{nN+N-1})(x) \end{bmatrix}$$

is orthonormal with respect to the matrix inner product

$$\langle P(x), Q(x)\rangle = \int P(x)dM(h^{-1}(x))Q^{\dagger}(x) + P(0)LQ^{\dagger}(0).$$

Here the matrix of measures $M(h^{-1}(x))$ *is defined by*

$$\int G(x)dM(h^{-1}(x)) =: \int G(h(x))dM(x).$$

Notice that $G(x)$ *is a vector function of* N *components, such that* $G(x) \in L_1(M)$*. As a conclusion, the Sobolev inner product can always be represented in terms of a matrix inner product, where the matrix of measures has a mass point supported at the origin.*

2.12 Example . *As a particular case, we take the following bilinear form*

$$B(f,g) = \int_0^{\infty} f(x)g(x)x^{\alpha}e^{-x}dx + C_0 f(0)g(0) + C_1 f'(0)g'(0). \tag{2.11}$$

The above bilinear form was studied in [49, 90, 91]. Notice that the lowest-degree polynomial $h(x)$ *such that* $B(hf,g) = B(f,hg)$ *is* $h(x) = x^2$*. Let* $(\tilde{L}^{\alpha}_{n,2}(x))_{n\in\mathbb{N}}$ *be the sequence of orthonormal polynomials with respect (2.11). Then the sequence of matrix polynomials*

$$\mathcal{L}_n(x) = \begin{bmatrix} R_{h,0}(\tilde{L}^{\alpha}_{2n,2})(x) & R_{h,1}(\tilde{L}^{\alpha}_{2n,2})(x) \\ R_{h,0}(\tilde{L}^{\alpha}_{2n+1,2})(x) & R_{h,1}(\tilde{L}^{\alpha}_{2n+1,2})(x) \end{bmatrix}$$

is orthogonal with respect to the matrix inner product

$$\langle P(x), Q(x)\rangle = \int_0^\infty P(x)dM(\sqrt{x})Q(x)^\dagger + P(0)\begin{bmatrix} C_0 & 0 \\ 0 & C_1 \end{bmatrix} Q(0)^\dagger,$$

where

$$dM(x) = \begin{bmatrix} 1 & x \\ x & x^2 \end{bmatrix} e^{-x}x^{\alpha}dx.$$

Chapter 3

Multiple Geronimus transformations

In 1939 J. Shohat [135] stated the following problem: Given a nontrivial probability measure μ supported on an interval of the real line, if $(p_n(x))_{n\in\mathbb{N}}$ denotes the corresponding sequence of monic orthogonal polynomials, then find necessary and sufficient conditions on the real numbers $A_1^{[n]} \neq 0$, with $n = 1,2\dots$, such that the sequence of monic polynomials $(Q_n(x))_{n\in\mathbb{N}}$ defined by

$$Q_n(x) := p_n(x) + A_1^{[n]} p_{n-1}(x), \qquad n \geq 1$$

is a family of orthogonal polynomials with respect to a measure supported on the real line. Shohat gave a partial answer by using Favard's Theorem. Few years after Shohat's publication, a complete answer to that problem was given by Geronimus in [76], providing a way to generate new families of orthogonal polynomials. The new family of orthogonal polynomials $(Q_n(x))_{n\in\mathbb{N}}$ is said to be a Geronimus transformation of $(p_n(x))_{n\in\mathbb{N}}$ (see [136], [149]).

3.1 The Geronimus transformation

Let $(p_n(x))_{n\in\mathbb{N}}$ be the sequence of monic orthogonal polynomials with respect to a quasi-definite moment functional σ. We know that $(p_n(x))_{n\in\mathbb{N}}$ satisfies a three term recurrence relation

$$xp_n(x) = p_{n+1}(x) + D_n p_n(x) + C_n p_{n-1}(x), \qquad n \geq 0, \quad p_{-1}(x) = 0,$$

where D_n and C_n are real numbers with $C_n \neq 0$. Using matrix notation, the above expression reads as

$$xp = J_{mon}p, \tag{3.1}$$

where

$$J_{mon} = \begin{bmatrix} D_0 & 1 & & \\ C_1 & D_1 & 1 & \\ & C_2 & D_2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

and $p = (p_0, p_1, \cdots)^\top$. The matrix J_{mon} is known as the monic Jacobi matrix associated with the sequence $(p_n(x))_{n\in\mathbb{N}}$. If J_{mon} can be written as $J_{mon} = UL + \kappa I$ where κ is a real number, (see [148] for conditions) and

$$U = \begin{bmatrix} \gamma_0 & 1 & & \\ & \gamma_1 & 1 & \\ & & \ddots & \ddots \end{bmatrix}, \qquad L = \begin{bmatrix} 1 & & & \\ \eta_1 & 1 & & \\ & \ddots & \ddots & \end{bmatrix},$$

then we can define a new semi-infinite matrix J^*_{mon} as $J^*_{mon} = LU + \kappa I$. From definition, J^*_{mon} is also a tridiagonal band matrix with 1's in the superdiagonal. Thus there exists a sequence of monic orthogonal polynomials $(p^*_n(x))_{n\in\mathbb{N}}$ associated with J^*_{mon} if and only if $\gamma_{n-1}\eta_n \neq 0$, $n \geq 1$ (see [148]). The following results appear in [33, 148].

Lemma 3.1. *If $J = UL + \kappa I$, then*

$$p^*_n(x) = p_n(x) + \eta_n p_{n-1}(x), \qquad n \geq 1,$$

and

$$(x-\kappa)p_n(x) = p^*_{n+1}(x) + \gamma_n p^*_n(x), \qquad n \geq 0.$$

The functional $\check{\sigma}$ for which the polynomials $(p^*_n(x))_{n\in\mathbb{N}}$ are orthogonal is given by $\sigma = (x-\kappa)\,\check{\sigma}$, or, equivalently, (see [99, 100, 148])

$$\check{\sigma} = (x-\kappa)^{-1}\sigma + \check{\sigma}_0 \delta(x-\kappa). \tag{3.2}$$

The linear functional $\check{\sigma}$ is called the canonical Geronimus transformation of σ. Notice that the constant $\check{\sigma}_0$ is an arbitrary real number. As a particular case, if in (3.2) we take the functional σ such that $\langle \sigma, f\rangle = \int_I f d\mu_0$ with μ_0 a nontrivial probability measure and $\kappa = 0$, then we have

$$\langle \check{\sigma}, fg\rangle = \int_I fg d\mu + \left(\check{\sigma}_0 - \int_I d\mu\right) f(0)g(0)$$

where $x d\mu = d\mu_0$. This expression appears in [48] in the framework of symmetric bilinear forms. The interest of considering symmetric bilinear forms in general is that the associated Gram matrix does not have the structure of Hankel matrix. This allows different kind of orthogonality, like orthogonality with respect to a Sobolev type inner product, which has been the main topic of many contributions in the last two decades (see [51], [65]). In this context, it is quite natural to define the Geronimus transformation as follows (see [48], [76]). Given a symmetric bilinear form on the linear space of polynomials $\mathbb{R}[x]$, defined by a nontrivial probability measure μ_0 supported on an infinite subset I of the real line

$$(f(x),g(x))_0 = \int_I f(x)g(x)d\mu_0(x),$$

its Geronimus transformation is the symmetric bilinear form $[\cdot,\cdot]_1$ such that

$$[xg(x),f(x)]_1 = [g(x),xf(x)]_1 = (f(x),g(x))_0 = \int_I f(x)g(x)d\mu_0(x).$$

In the first part of [48] the authors show that the sequence of monic orthogonal polynomials $(p_n^*(x))_{n\in\mathbb{N}}$ associated with $[\cdot,\cdot]_1$ can written as

$$p_n^*(x) = p_n(x) + A_n p_{n-1}(x)$$

where $(p_n(x))_{n\in\mathbb{N}}$ is the sequence of monic orthogonal polynomials with respect to $(\cdot,\cdot)_0$. Furthermore, if J_{mon} and J_{mon}^* are the Jacobi matrices associated with the sequences $(p_n(x))_{n\in\mathbb{N}}$ and $(p_n^*(x))_{n\in\mathbb{N}}$, respectively, then there exist an upper triangular matrix U and a lower triangular matrix L such that

$$J_{mon} = UL \quad \text{and} \quad J_{mon}^* = LU.$$

In the second part, the authors present the double Geronimus transformation in the framework of symmetric bilinear forms defined as

$$[x^2g(x),f(x)]_2 = [g(x),x^2f(x)]_2 = (f(x),g(x))_0 = \int_I f(x)g(x)d\mu_0(x).$$

Furthermore, they deduce that the monic orthogonal polynomials $(p_n^{**}(x))_{n\in\mathbb{N}}$ associated with $[\cdot,\cdot]_2$ satisfy the following connection formula

$$p_n^{**}(x) = p_n(x) + A_n p_{n-1}(x) + B_n p_{n-2}(x),$$

where A_n and B_n are constants with $B_n \neq 0$ ([48], Theorem 4.3). Moreover, if we define J_{mon}^{**} as the matrix associated with $(p_n^{**}(x))_{n\in\mathbb{N}}$ and

$$J^{**} = \left([x^2\tilde{p}_n^{**}(x),\tilde{p}_m^{**}(x)]_2\right)_{n,m=0}^{\infty}, \text{ where, } p_n^{**}(x) = \frac{1}{h_n^{**}}\tilde{p}_n^{**}(x),$$

with $(h_n^{**})^2 = [p_n^{**}(x),p_n^{**}(x)]_2$, $h_n^{**} > 0$, then there exist an upper triangular matrix U and lower triangular matrices L and C such that

$$J_{mon}^2 = UL, \qquad J_{mon}^{**} = LU, \qquad \text{and} \qquad J^{**} = CC^{\top}.$$

Next, let us consider a polynomial $h(x)$ with $deg(h) = N$. A natural question is what can be said about the symmetric bilinear form $[\cdot,\cdot]_h$ defined as

$$[hg(x),f(x)]_h = [g(x),hf(x)]_h = \int_I f(x)g(x)d\mu_0(x). \tag{3.3}$$

This problem is motivated by A. J. Durán in [51], where general results are given for symmetric bilinear forms such that the multiplication operator by $h(x)$ is symmetric with respect to the bilinear form, i.e. $B(hf,g) = B(f,gh)$.

The next subsection is dedicated to give some results concerning the symmetric bilinear form defined in (3.3), finding as particular cases the results given in [48].

3.2 An extension of the Geronimus transformation to the multiple case

Let us consider a symmetric bilinear form

$$(f,g)_0 = \int_I f(x)g(x)d\mu_0(x),$$

where μ_0 is a nontrivial probability measure supported on an infinite subset I of the real line. In general, if we assume that $(\cdot,\cdot)_0$ is quasi-definite, then we know that the corresponding sequence of monic orthogonal polynomials $(p_n(x))_{n\in\mathbb{N}}$ satisfies a three term recurrence relation

$$xp_n(x) = p_{n+1}(x) + D_n p_n(x) + C_n p_{n-1}(x), \quad n \geq 0, \quad p_0(x) = 1, \quad p_{-1}(x) = 0,$$

where D_n and C_n are real numbers with $C_n \neq 0$. Using matrix notation, the above expression reads $xp = J_{mon}p$ where J_{mon} is as in (3.1). If we assume that $(\cdot,\cdot)_0$ is a positive definite bilinear form, then there is a sequence of orthonormal polynomials $(\check{p}_n(x))_{n\in\mathbb{N}}$ such that

$$x\check{p}_n(x) = \check{C}_{n+1}\check{p}_{n+1}(x) + D_n\check{p}_n(x) + \check{C}_n\check{p}_{n-1}(x), \quad n \geq 0.$$

Notice that in such a case, $C_n = \check{C}_n^2 > 0$. We can associate with the above orthonormal sequence a Jacobi tridiagonal symmetric matrix

$$\check{J} = \begin{bmatrix} D_0 & \check{C}_1 & & \\ \check{C}_1 & D_1 & \check{C}_2 & \\ & \check{C}_2 & D_2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}.$$

such that $x\check{p} = \check{J}\check{p}$, where $\check{p} = [\check{p}_0(x), \check{p}_1(x), \check{p}_2(x), \cdots]^\top$.

Let $h(x)$ be a monic polynomial with $deg(h) = N$. Let us define a symmetric bilinear form $[\cdot,\cdot]_h$ on the linear space $\mathbb{R}[x]$

$$[hf,g]_h = [f,hg]_h = \int_I f(x)g(x)d\mu_0(x).$$

Clearly, this definition does not determine the bilinear form $[\cdot,\cdot]_h$ uniquely. Moreover , we can choose the entries of the following symmetric matrix

$$\check{S} = \begin{bmatrix} [1,1]_h & \cdots & [1,x^{N-1}]_h \\ \vdots & \vdots & \vdots \\ [x^{N-1},1]_h & \cdots & [x^{N-1},x^{N-1}]_h \end{bmatrix} = \begin{bmatrix} s_{0,0} & \cdots & s_{0,N-1} \\ \vdots & \vdots & \vdots \\ s_{N-1,0} & \cdots & s_{N-1,N-1} \end{bmatrix} \tag{3.4}$$

in an arbitrary way. It should be noted that because the multiplication operator by $h(x)$ is symmetric for $[\cdot,\cdot]_h$, if we assume that this symmetric bilinear form is quasi-definite, then the corresponding sequence of monic orthogonal polynomials $(p_n^*(x))_{n\in\mathbb{N}}$ satisfies

$$h(x)p_n^*(x) = \sum_{k=n-N}^{n+N} c_k^{[n]} p_k^*(x),$$

where $c_{n+N}^{[n]} = 1$, and $c_{n-N}^{[n]} > 0$ for $n \geq N$. Thus, we can associate with the sequence $(p_n^*(x))_{n\in\mathbb{N}}$ the $2N+1$ band matrix

$$J_{mon}^* = \begin{bmatrix} c_0^{[0]} & c_1^{[0]} & \cdots & c_{N-1}^{[0]} & 1 & & & & \\ c_0^{[1]} & c_1^{[1]} & \cdots & \cdots & c_N^{[1]} & 1 & & & \\ \vdots & \vdots & \ddots & & & \ddots & \ddots & & \\ c_0^{[N]} & c_1^{[N]} & \cdots & \ddots & & & c_{2N-1}^{[N]} & 1 & \\ 0 & c_1^{[N+1]} & \cdots & & \ddots & & & c_{2N}^{[N+1]} & 1 \\ & & \ddots & & & & & \ddots & \ddots \end{bmatrix}.$$

Before dealing with the properties of the symmetric bilinear form $[\cdot,\cdot]_h$ we will choose an appropriate basis in the linear space $\mathbb{R}[x]$. Indeed, let consider the following basis

$$\mathfrak{B}_h = \{x^m h^k,\ k \geq 0, \quad 0 \leq m \leq N-1\}.$$

Recall that this allows us to express every polynomial f as $f(x) = \sum_{k\geq 0}\sum_{m=0}^{N-1} a_{k,m}x^m h^k(x)$ (see (2.9)). Moreover, if we define for a fixed k the following linear operator

$$H_{k,h}(f)(x) = \sum_{m=0}^{N-1} a_{k,m}x^m h^k(x),$$

then we have $f(x) = \sum_{k=0} H_{k,h}(f)(x)$.

Let $x_1,\cdots,x_q$, be the zeros of $h(x)$ and $\alpha_1,\cdots,\alpha_q$, be their corresponding multiplicities, i.e.

$$h(x) = (x-x_1)^{\alpha_1}(x-x_2)^{\alpha_2}\cdots(x-x_q)^{\alpha_q}, \qquad \text{with} \qquad \sum_{i=1}^{q}\alpha_i = N.$$

For each x_i, $h(x)$ can be represented as

$$h(x) = (x - x_i)^{\alpha_i} t_i(x),$$

where $t_i(x)$ is a polynomial of degree $N - \alpha_i$ such that $t_i(x_i) \neq 0$. According to the Leibniz product rule for derivatives

$$h^{(j)}(x) = \sum_{k=0}^{j} \binom{j}{k} \frac{\alpha_i!}{(\alpha_i - k)!} (x - x_i)^{\alpha_i - k} t_i^{(j-k)}(x).$$

Thus

$$h^{(j)}(x_i) = 0 \quad \text{for} \quad j = 0, \cdots, \alpha_i - 1.$$

If f is a polynomial, using its representation in terms of the basis defined as above, then we get

$$f^{(j)}(x_i) = \sum_{m=j}^{N-1} a_{0,m} \frac{m!}{(m-j)!} x_i^{m-j}.$$

Thus, for each $i = 1, \cdots, q$, we have

$$\begin{bmatrix} f(x_i) \\ f^{(1)}(x_i) \\ \vdots \\ \vdots \\ f^{(\alpha_i - 2)}(x_i) \\ f^{(\alpha_i - 1)}(x_i) \end{bmatrix} = \begin{bmatrix} 1 & x_i & x_i^2 & x_i^3 & \cdots & \cdots & \cdots & x_i^{N-1} \\ & 1 & 2x_i & 3x_i^2 & \cdots & \cdots & \cdots & (N-1)x_i^{N-2} \\ & & 2! & 6x_i & \cdots & \cdots & \cdots & (N-1)(N-2)x_i^{N-3} \\ & & & \ddots & \ddots & \cdots & \cdots & \vdots \\ & & & & (\alpha_i - 1)! & \alpha_i x_i & \cdots & \frac{(N-1)!}{(N-\alpha_i)!} x_i^{N-\alpha_i} \end{bmatrix}_{\alpha_i \times N} \begin{bmatrix} a_{0,0} \\ a_{0,1} \\ \vdots \\ \vdots \\ a_{0,N-2} \\ a_{0,N-1} \end{bmatrix}$$

or, equivalently,

$$F_i = A_i \begin{bmatrix} a_{0,0} \\ \vdots \\ a_{0,N-1} \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} F_1 \\ \vdots \\ F_p \end{bmatrix} = \begin{bmatrix} A_1 \\ \vdots \\ A_p \end{bmatrix}_{N \times N} \begin{bmatrix} a_{0,0} \\ \vdots \\ a_{0,N-1} \end{bmatrix} = \mathcal{A} \begin{bmatrix} a_{0,0} \\ \vdots \\ a_{0,N-1} \end{bmatrix}. \tag{3.5}$$

By the definition of the polynomial $f(x)$, the above system of linear equations (3.5) has at least a solution $[a_{0,0}, \cdots, a_{0,N-1}]^\top$. Let us assume that there is another solution denoted by $[a'_{0,0}, \cdots, a'_{0,N-1}]^\top$. Then we define the polynomials $u(x) = \sum_{m=0}^{N-1} a_{0,m} x^m$ and $v(x) = \sum_{m=0}^{N-1} a'_{0,m} x^m$. So, according to (3.5) we have that for each $i = 1, \cdots, q$,

$$u^{(j)}(x_i) = f^{(j)}(x_i) = v^{(j)}(x_i), \qquad \text{for} \quad j = 0, \cdots, \alpha_i - 1.$$

We now define the polynomial $c(x) = u(x) - v(x)$. Notice that $deg(c(x)) \leq N-1$. On the other hand,

$$c^{(j)}(x_i) = 0 \qquad \text{for} \quad j = 0, \cdots, \alpha_i - 1.$$

This implies that x_i is a zero of multiplicity at least α_i for $c(x)$ and since it is true for every $i = 1, \cdots, q$, then $deg(c(x)) \geq N$. So, necessarily $c(x) = 0$, i.e. $u(x) = v(x)$. Therefore the solution of (3.5) is unique and, as a consequence, $\mathcal{A}$ is a nonsingular matrix. In particular, if the zeros of $h(x)$ are simple, then (3.5) takes the form

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{N-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_N & \cdots & x_N^{N-1} \end{bmatrix}_{N\times N} \begin{bmatrix} a_{0,0} \\ \vdots \\ a_{0,N-1} \end{bmatrix}.$$

In other words, $\mathcal{A}$ is a Vandermonde matrix.

Proposition 3.2. *Let μ be a nontrivial probability measure supported on the real line such that $d\mu_0 = hd\mu$ has finite moments. Let $f(x) = \sum a_{k,m}x^m h^k(x)$ and $g(x) = \sum b_{k',m'}x^{m'}h^{k'}(x)$ be polynomials. Then $[\cdot,\cdot]_h$ can be represented as follows*

$$[f,g]_h = \int_I f(x)g(x)d\mu + \begin{bmatrix} F_1^\top & \cdots & F_q^\top \end{bmatrix} \mathcal{A}^{-\top} S \mathcal{A}^{-1} \begin{bmatrix} G_1 \\ \vdots \\ G_q \end{bmatrix}, \tag{3.6}$$

where $G_i = [g(x_i), \cdots, g^{(\alpha_i - 1)}(x_i)]^\top$, $F_i = [f(x_i), \cdots, f^{(\alpha_i-1)}(x_i)]^\top$ and S is a symmetric matrix with

$$S = \begin{bmatrix} s_{0,0} - \int_I d\mu & \cdots & s_{0,N-1} - \int_I x^{N-1} d\mu \\ \vdots & \vdots & \vdots \\ s_{N-1,0} - \int_I x^{N-1} d\mu & \cdots & s_{N-1,N-1} - \int_I x^{2N-2} d\mu \end{bmatrix}.$$

Proof. To compute $[f,g]_h$, let us observe that the polynomial

$$f(x) - \sum_{m=0}^{N-1} a_{0,m}x^m$$

is divisible by $h(x)$ by construction. Now we have

$$[f,g]_h = \left[f(x) - \sum_{m=0}^{N-1} a_{0,m}x^m, g(x) \right]_h + \left[\sum_{m=0}^{N-1} a_{0,m}x^m, g(x) \right]_h$$

$$= \left(\frac{f(x) - \sum_{m=0}^{N-1} a_{0,m} x^m}{h}, g(x) \right)_0 + \left[\sum_{m=0}^{N-1} a_{0,m} x^m, g(x) - \sum_{m'=0}^{N-1} b_{0,m'} x^{m'} \right]_h + \left[\sum_{m=0}^{N-1} a_{0,m} x^m, \sum_{m'=0}^{N-1} b_{0,m'} x^{m'} \right]_h$$
$$= \int_I \left(\frac{f(x) - \sum_{m=0}^{N-1} a_{0,m} x^m}{h} \right) g(x) d\mu_0$$
$$+ \int_I \sum_{m=0}^{N-1} a_{0,m} x^m \left(\frac{g(x) - \sum_{m'=0}^{N-1} b_{0,m'} x^{m'}}{h} \right) d\mu_0 + \sum_{m=0}^{N-1} \sum_{m'=0}^{N-1} a_{0,m} b_{0,m'} [x^m, x^{m'}]_h$$
$$= \int_I f(x) g(x) d\mu - \sum_{m=0}^{N-1} \sum_{m'=0}^{N-1} a_{0,m} b_{0,m'} \int_I x^{m'+m} d\mu + \sum_{m=0}^{N-1} \sum_{m'=0}^{N-1} a_{0,m} b_{0,m'} s_{m,m'},$$

where $s_{m,m'} = [x^m, x^{m'}]_h$. In a matrix form the above expression reads

$$[f,g]_h = \int_I f(x) g(x) d\mu +$$
$$\begin{bmatrix} a_{0,0} & \cdots & a_{0,N-1} \end{bmatrix} \begin{bmatrix} s_{0,0} - \int_I d\mu & \cdots & s_{0,N-1} - \int_I x^{N-1} d\mu \\ \vdots & \vdots & \vdots \\ s_{N-1,0} - \int_I x^{N-1} d\mu & \cdots & s_{N-1,N-1} - \int_I x^{2N-2} d\mu \end{bmatrix} \begin{bmatrix} b_{0,0} \\ \vdots \\ b_{0,N-1} \end{bmatrix}.$$

Next, using (3.5) we get (3.6). ■

If we assume that $h(x) = x^N$, then we have the following result that appears in [76] for $N = 2$.

Corollary 3.3. *If μ is a nontrivial probability measure supported in the real line such that $d\mu_0 = x^N d\mu$ has finite moments, then*

$$[f,g]_h = \int_I f(x) g(x) d\mu + \begin{bmatrix} f(0) & \cdots & f^{(N-1)}(0) \end{bmatrix} M \begin{bmatrix} g(0) \\ \vdots \\ g^{(N-1)}(0) \end{bmatrix}, \tag{3.7}$$

where M is a symmetric matrix such that

$$M = \begin{bmatrix} \frac{1}{0!} & & \\ & \ddots & \\ & & \frac{1}{(N-1)!} \end{bmatrix} S \begin{bmatrix} \frac{1}{0!} & & \\ & \ddots & \\ & & \frac{1}{(N-1)!} \end{bmatrix}.$$

Since the values $s_{i,j}$ in (3.4) are arbitrary, if we take them in such a way that the matrix S is diagonal, i.e.

$$S = \begin{bmatrix} \lambda_0 & & \\ & \ddots & \\ & & \lambda_{N-1} \end{bmatrix},$$

then (3.7) reduces to

$$[f,g]_h = \int_I f(x)g(x)d\mu + \sum_{k=0}^{N-1} M_k f^{(k)}(0)g^{(k)}(0) \quad \text{with} \quad M_k = \frac{\lambda_k}{(k!)^2},$$

which is a diagonal Sobolev-type inner product. In other words, a N-th iterated Geronimus transformation of $(\cdot,\cdot)_0$ generates Sobolev type inner products.

3.2.1 Orthogonal polynomials associated to the extension of the Geronimus transformation

Next, assuming that the bilinear form $[\cdot,\cdot]_h$ is quasi-definite, we will represent the monic polynomials $(p_n^*(x))_{n\in\mathbb{N}}$, orthogonal with respect to $[\cdot,\cdot]_h$, in terms of the sequence $(p_n(x))_{n\in\mathbb{N}}$ of monic orthogonal polynomials with respect to $(\cdot,\cdot)_0$. Notice that from the orthogonality hypothesis, for the elements of the basis $\mathfrak{B}_h$ we get

$$[p_n^*(x), x^m h^k]_h = [x^m h^k, p_n^*(x)]_h = 0, \quad \text{for} \quad Nk+m \leq n-1.$$

So, for $n > N$, from the definition of the bilinear form we get

$$[p_n^*(x), x^m h^k]_h = (p_n^*(x), x^m h^{k-1})_0 = 0, \quad \text{for} \quad N(k-1)+m < n-N \text{ and } k \geq 1,$$

which means that

$$p_n^*(x) = p_n(x) + A_{n-1}^{[n]} p_{n-1}(x) + \cdots + A_{n-N}^{[n]} p_{n-N}(x). \tag{3.8}$$

At the same time, we also have

$$[p_n^*(x), x^m]_h = 0, \quad \text{for} \quad m = 0, \cdots, N-1,$$

which can be rewritten as

$$[p_n(x), x^m]_h + A_{n-1}^{[n]} [p_{n-1}(x), x^m]_h + \cdots + A_{n-N}^{[n]} [p_{n-N}(x), x^m]_h = 0.$$

The last relation is equivalent to the following system of linear equations

$$\begin{bmatrix} [p_{n-1}(x),1]_h & \cdots & [p_{n-N}(x),1]_h \\ \vdots & & \vdots \\ [p_{n-1}(x),x^{N-1}]_h & \cdots & [p_{n-N}(x),x^{N-1}]_h \end{bmatrix} \begin{bmatrix} A_{n-1}^{[n]} \\ \vdots \\ A_{n-N}^{[n]} \end{bmatrix} = \begin{bmatrix} -[p_n(x),1]_h \\ \vdots \\ -[p_n(x),x^{N-1}]_h \end{bmatrix}. \tag{3.9}$$

Lemma 3.4. *The system*

$$\begin{bmatrix} [p_{n-1}(x),1]_h & \cdots & [p_{n-N}(x),1]_h \\ \vdots & & \vdots \\ [p_{n-1}(x),x^{N-1}]_h & \cdots & [p_{n-N}(x),x^{N-1}]_h \end{bmatrix} b = \begin{bmatrix} -[p_n(x),1]_h \\ \vdots \\ -[p_n(x),x^{N-1}]_h \end{bmatrix}$$

has an unique solution $b = \begin{bmatrix} A_{n-1}^{[n]} \\ \vdots \\ A_{n-N}^{[n]} \end{bmatrix}$.

Proof. Since $p_n^*(x)$ is a monic polynomial of degree n, we know that (3.9) has at least one solution. Suppose that the above system has other solution $(A_{n-1}^{[n]\prime} \cdots A_{n-N}^{[n]\prime})^\top$ and we define the polynomial

$$Q_n(x) = p_n(x) + A_{n-1}^{[n]\prime} p_{n-1}(x) + \cdots + A_{n-N}^{[n]\prime} p_{n-N}(x).$$

From hypothesis $[Q_n(x), x^m]_h = 0$ for $m = 0, \ldots, N-1$. On the other hand, let $Nk + m < n$ with $k \neq 0$, then

$$\begin{aligned}
[Q_n(x), h^k x^m]_h &= [p_n(x), h^k x^m]_h + A_{n-1}^{[n]\prime}[p_{n-1}(x), h^k x^m]_h + \cdots + A_{n-N}^{[n]\prime}[p_{n-N}(x), h^k x^m]_h \\
&= (p_n(x), h^{k-1} x^m)_0 + A_{n-1}^{[n]\prime}(p_{n-1}(x), h^{k-1} x^m)_0 + \cdots + A_{n-N}^{[n]\prime}(p_{n-N}(x), h^{k-1} x^m)_0 \\
&= 0
\end{aligned}$$

and $[Q_n(x), x^n]_h = [Q_n(x), p_n^*(x)]_h$. The above implies that, for $0 \leq \ell \leq n$, $[Q_n(x), x^\ell]_h = \|p_n^*(x)\|_h \delta_{n,\ell}$. Thus, if we assume that the system has two different solutions, then there are two monic polynomials of degree n that satisfy the orthogonality condition. But this contradicts the uniqueness of the sequence $(p_n^*(x))_{n\in\mathbb{N}}$. ■

Moreover, the uniqueness also yields

$$d_n^* = \begin{vmatrix} [p_{n-1}(x), 1]_h & \cdots & [p_{n-N}(x), 1]_h \\ \vdots & & \vdots \\ [p_{n-1}(x), x^{N-1}]_h & \cdots & [p_{n-N}(x), x^{N-1}]_h \end{vmatrix} \neq 0. \tag{3.10}$$

Furthermore, according to the Cramer's rule, the polynomials $p_n^*(x)$ can be represented as follows

$$p_n^*(x) = \frac{1}{d_n^*} \begin{vmatrix} p_n(x) & [p_n(x), 1]_h & \cdots & [p_n(x), x^{N-1}]_h \\ \vdots & \vdots & \cdots & \vdots \\ p_{n-i}(x) & [p_{n-i}(x), 1]_h & \cdots & [p_{n-i}(x), x^{N-1}]_h \\ \vdots & \vdots & \cdots & \vdots \\ p_{n-N}(x) & [p_{n-N}(x), 1]_h & \cdots & [p_{n-N}(x), x^{N-1}]_h \end{vmatrix}.$$

Now, for $0 \leq l \leq N-1$ and $H_{0,h}(p_j)(x) = \sum_{k=0}^{N-1} c_{0,k} x^k$, we get

$$\begin{aligned}
[p_j, x^l]_h &= [p_j(x) - H_{0,h}(p_j)(x) + H_{0,h}(p_j)(x), x^l]_h \\
&= [\sum_{m\geq 0} H_{m,h}(p_j)(x) - H_{0,h}(p_j)(x), x^l]_h + [\sum_{k=0}^{N-1} c_{0,k} x^k, x^l]_h \\
&= \left(\frac{\sum_{m\geq 0} H_{m,h}(p_j)(x) - H_{0,h}(p_j)(x)}{h}, x^l \right)_0 + \sum_{k=0}^{N-1} c_{0,k} s_{k,l} \\
&= \sum_{m\geq 1} \int_I \frac{H_{m,h}(p_j)(x)}{h} x^l d\mu_0 + \sum_{k=0}^{N-1} c_{0,k} s_{k,l}.
\end{aligned}$$

Let us stress that the above analysis was done for $n \geq N$. However, it is clear that, for $n \leq N$, the polynomial $p_n^*(x)$ has the following form

$$p_n^*(x) = p_n(x) + A_{n-1}^{[n]} p_{n-1}(x) + \cdots + A_0^{[n]} p_0(x),$$

where we put $p_m(x) = 0$ for $m < 0$. So, if we use similar arguments as above, then for $n \leq N$ we get

$$p_n^*(x) = \frac{1}{d_n^*} \begin{vmatrix} p_n(x) & [p_n(x),1]_h & \cdots & [p_n(x),x^{n-1}]_h \\ \vdots & \vdots & \cdots & \vdots \\ p_{n-i}(x) & [p_{n-i}(x),1]_h & \cdots & [p_{n-i}(x),x^{n-1}]_h \\ \vdots & \vdots & \cdots & \vdots \\ p_0(x) & [p_0(x),1]_h & \cdots & [p_0(x),x^{n-1}]_h \end{vmatrix}.$$

As a last remark, let us notice that if $n < N$, then d_n^* is the determinant of a matrix of size $n \times n$, which depends on n, while in the other cases d_n^* is the determinant of a matrix of size $N \times N$.

Thus, we can deduce the following

Proposition 3.5. *Let $(\cdot,\cdot)_0$ be a quasi-definite bilinear form and let $(p_n(x))_{n\in\mathbb{N}}$ be the corresponding sequence of monic orthogonal polynomials. Then the symmetric bilinear form $[\cdot,\cdot]_h$ is quasi-definite if only if $d_n^* \neq 0$ for all $n \in \mathbb{N}$. In such a case, the sequence the monic polynomials $(p_n^*(x))_{n\in\mathbb{N}}$ defined as above is orthogonal with respect to $[\cdot,\cdot]_h$. Furthermore, we get the following representation*

$$p_n^*(x) = \frac{1}{d_n^*} \begin{vmatrix} p_n(x) & [p_n(x),1]_h & \cdots & [p_n(x),x^{N-1}]_h \\ \vdots & \vdots & \cdots & \vdots \\ p_{n-i}(x) & [p_{n-i}(x),1]_h & \cdots & [p_{n-i}(x),x^{N-1}]_h \\ \vdots & \vdots & \cdots & \vdots \\ p_{n-N}(x) & [p_{n-N}(x),1]_h & \cdots & [p_{n-N}(x),x^{N-1}]_h \end{vmatrix}, \tag{3.11}$$

where d_n^* is defined by (3.10) *and*

$$[p_j(x),x^l]_h = \sum_{m\geq 1} \int_I \frac{H_{m,h}(p_j)(x)}{h} x^l d\mu_0 + \sum_{k=0}^{N-1} c_{0,k} s_{k,l}.$$

If we assume that the zeros of $h(x)$ are outside the interior of the convex hull of support of μ_0, then the bilinear form $[f,g]_\mu = \int_I fg d\mu$, where $\mu = \frac{\mu_0}{h(x)}$ is positive definite. With this in mind, we can state the following

Corollary 3.6. *If $(r_n(x))_{n\in\mathbb{N}}$ is the sequence of monic polynomials orthogonal with respect to $[\cdot,\cdot]_\mu$, then we get the following connection formula for $(p_n^*(x))_{n\in\mathbb{N}}$ and $(r_n(x))_{n\in\mathbb{N}}$*

$$h(x)p_n^*(x) = r_{n+N}(x) + B_{n+N-1}^{[n]} r_{n+N-1}(x) + \cdots + B_{n-N}^{[n]} r_{n-N}(x).$$

Furthermore,

$$(p_{n+N}^*(x), r_k(x))_0 = 0, \quad \textit{if} \quad k < n. \tag{3.12}$$

Proof. Notice that $h(x)p_n(x)$ can be written as

$$h(x)p_n(x) = \sum_{k=0}^{n+N} b_k^{[n]} r_k(x),$$

where

$$b_k^{[n]} = \frac{[h(x)p_n(x), r_k(x)]_\mu}{\|r_k\|_\mu^2} = \frac{(p_n(x), r_k(x))_0}{\|r_k\|_\mu^2} = \begin{cases} 0, & k < n, \\ \frac{(p_n(x), r_k(x))_0}{\|r_k\|_\mu^2}, & n \leq k. \end{cases}$$

But (3.8) immediately yields

$$h(x)p_n^*(x) = r_{n+N}(x) + B_{n+N-1}^{[n]} r_{n+N-1}(x) + \cdots + B_{n-N}^{[n]} r_{n-N}(x),$$

where

$$B_{n+N-m}^{[n]} = \sum_{k=0}^{min\{m,N\}} b_{N+n-m}^{[n-k]} A_{n-k}^{[n]}.$$

On the other hand,

$$h(x)p_n^*(x) = \sum_{k=0}^{N+n} c_k^{[n]} r_k(x) \text{ with } c_k^{[n]} = \frac{(r_k(x), p_n^*(x))_0}{\|r_k\|_\mu^2}.$$

According to (3.11), we get $c_k^{[n]} = 0$, $0 \leq k \leq n-N-1$, and $c_{n-N}^{[n]} \neq 0$.
Taking into account the representation of $h(x)p_n^*(x)$ in terms of the sequence $(r_n(x))_{n\in\mathbb{N}}$ is unique, then (3.12) holds. ■

3.7 Example . *Let us assume that $h(x) = x^N$, $d\mu_0 = x^{\alpha+N}e^{-x}dx$, and let us define $(\cdot,\cdot)_0$ as*

$$(f,g)_0 = \int_0^\infty f(x)g(x)x^{\alpha+N}e^{-x}dx, \qquad \alpha > -1.$$

We know that the sequence of monic orthogonal polynomials associated with the above bilinear form are the Laguerre polynomials $(L_n^{\alpha+N}(x))_{n\in\mathbb{N}}$ with parameter $\alpha+N$. Let us now take $d\mu = x^\alpha e^{-x}dx$. Then

$$[f(x), g(x)]_h = \int_0^\infty f(x)g(x)x^\alpha e^{-x}dx + \sum_{k,j=0}^{N-1} M_{k,j} f^{(k)}(0)g^{(j)}(0). \tag{3.13}$$

As a straightforward consequence, the sequence of polynomials orthogonal with respect to (3.13) can be written as

$$\tilde{L}_n^{\alpha}(x) = L_n^{\alpha+N}(x) + \sum_{k=1}^{N} A_{n-k}^{[n]} L_{n-k}^{\alpha+N}(x).$$

The above bilinear form and their orthogonal polynomials are very well known in the literature. Indeed, the diagonal case was introduced in [90]. Let us notice that, in particular, if $M_{k,j} = 0$ for $(k,j) \neq (0,0)$ we get the so called Laguerre-Krall orthogonal polynomials (see [85]).

The previous corollary shows a connection formula between the sequences of polynomials $(p_n^*(x))_{n\in\mathbb{N}}$ and $(r_n(x))_{n\in\mathbb{N}}$. Next, we focus our attention to find necessary and sufficient conditions for the existence of the sequence of polynomials $(p_n^*(x))_{n\in\mathbb{N}}$.

The condition given in (3.6), can be rewritten as

$$[f,g]_h = \int_I fg d\mu + \sum_{l,w=1}^{q} \sum_{i=0}^{\alpha_l-1} \sum_{j=0}^{\alpha_w-1} \lambda_{i,j,l,w} f^{(i)}(x_l) g^{(j)}(x_w).$$

Thus

$$\begin{aligned}
p_n^*(x) &= r_n(x) + \sum_{k=0}^{n-1} \left[-\sum_{l,w=1}^{q} \sum_{i=0}^{\alpha_l-1} \sum_{j=0}^{\alpha_w-1} \lambda_{i,j,l,w} \left(p_n^*\right)^{(i)}(x_l) r_k^{(j)}(x_w) \right] r_k(x) \\
&= r_n(x) - \sum_{l,w=1}^{q} \sum_{i=0}^{\alpha_l-1} \sum_{j=0}^{\alpha_w-1} \lambda_{i,j,l,w} \left(p_n^*\right)^{(i)}(x_l) \left(\sum_{k=0}^{n-1} \frac{r_k^{(j)}(x_w) r_k(x)}{\|r_k\|_\mu^2} \right) \\
&= r_n(x) - \sum_{l=1}^{q} \sum_{i=0}^{\alpha_l-1} \left(p_n^*\right)^{(i)}(x_l) D_{i,l}(x),
\end{aligned}$$

where

$$D_{i,l}(x) = \sum_{w=1}^{q} \sum_{j=0}^{\alpha_w-1} \lambda_{i,j,l,w} K_{n-1}^{(j,0)}(x_w, x).$$

In particular, for $1 \le a \le q$ and $1 \le k \le \alpha_a - 1$,

$$r_n^{(k)}(x_a) = \left(p_n^*\right)^{(k)}(x_a) + \sum_{l=1}^{q} \sum_{i=0}^{\alpha_l-1} \left(p_n^*\right)^{(i)}(x_l) D_{i,l}^{(k)}(x_a).$$

Let introduce the vector

$$v_j^k(a) = \begin{cases} [D_{0,j}^{(k)}(x_j), \cdots, 1 + D_{k,j}^{(k)}(x_j), \cdots, D_{\alpha_k-1,j}^{(k)}(x_j)], \text{if} \quad j = a, \\ \\ [D_{0,j}^{(k)}(x_j), \cdots, D_{k,j}^{(k)}(x_j), \cdots, D_{\alpha_k-1,j}^{(k)}(x_j)], \text{if} \quad j \neq a. \end{cases}$$

Then for each $1 \leq a \leq q$,

$$\mathbb{R}_a = \begin{bmatrix} r_n(x_a) \\ \vdots \\ r_n^{(\alpha_a-1)}(x_a) \end{bmatrix} = \begin{bmatrix} v_1^0(a) & v_2^0(a) & \cdots & v_q^0(a) \\ \vdots & \vdots & & \vdots \\ v_1^{\alpha_a-1}(a) & v_2^{\alpha_a-1}(a) & & v_q^{\alpha_a-1}(a) \end{bmatrix}_{\alpha_a \times N} \begin{bmatrix} (p_n^*)^{(0)}(x_1) \\ \vdots \\ (p_n^*)^{(\alpha_1-1)}(x_1) \\ \vdots \\ (p_n^*)^{(0)}(x_q) \\ \vdots \\ (p_n^*)^{(\alpha_q-1)}(x_q) \end{bmatrix} = \mathbb{V}_a \mathbb{P}^*.$$

From the above we can state the following

Proposition 3.8. *Let μ be a positive Borel measure supported in the real line and $(r_n(x))_{n\in\mathbb{N}}$ the sequence of orthogonal polynomials with respect to the bilinear form $[f,g]_\mu := \int_I fgd\mu$. Let $[\cdot,\cdot]_h$ be a symmetric bilinear form*

$$[f,g]_h = \int_I fgd\mu + \sum_{l,w}^{q} \sum_{i=0}^{\alpha_l-1} \sum_{j=0}^{\alpha_w-1} \lambda_{i,j,l,w} f^{(i)}(x_l) g^{(j)}(x_w),$$

with $\lambda_{i,j,l,w} = \lambda_{j,i,w,l}$. A necessary and sufficient condition for the existence of the sequence of monic polynomials $(p_n^*(x))_{n\in\mathbb{N}}$ orthogonal with respect to $[\cdot,\cdot]_h$ is that the system of linear equations

$$\begin{bmatrix} \mathbb{R}_1 \\ \vdots \\ \mathbb{R}_q \end{bmatrix} = \begin{bmatrix} \mathbb{V}_1 \\ \vdots \\ \mathbb{V}_q \end{bmatrix} \mathbb{P}^*$$

has an unique solution.

As a next step, a natural question can be posed: when the bilinear form $[\cdot,\cdot]_h$ is positive definite?. It is clear that if we assume that the matrix S is a positive semidefinite matrix, as μ is a positive Borel measure, then $[\cdot,\cdot]_h$ is positive definite, because given $t(x)$, a polynomial neither identically zero nor negative in the support I, we have that

$$[t(x),t(x)]_h = \int_I t^2(x)d\mu + v^\top S v \geq 0,$$

where

$$v^\top = \begin{bmatrix} T_1^\top & \cdots & T_q^\top \end{bmatrix} \mathcal{A}^{-\top} \qquad \text{with} \qquad T_i = [t(x_i), \cdots, t^{(\alpha_i-1)}(x_i)]^\top,$$

but this does not give us so much information. Alternatively, in order to analyze the positivity of $[p_n^*(x), p_n^*(x)]_h$, we need to consider two cases : When $n = m + Nk$ with $k \neq 0$ and when $n < N$.

Case 1. If $n = m + Nk$, with $k \neq 0$, then

$$[p_n^*(x), x^m h^k]_h = \int_I p_n^*(x) x^m h^k d\mu = \int_I p_n^*(x) x^m h^{k-1} d\mu_0$$
$$= \frac{1}{d_n^*} \begin{vmatrix} \int p_n(x) x^m h^{k-1} d\mu_0 & [p_n(x), 1]_h & \cdots & [p_n(x), x^{N-1}]_h \\ \vdots & \vdots & & \vdots \\ \int p_{n-N+1}(x) x^m h^{k-1} d\mu_0 & [p_{n-N+1}(x), 1]_h & \cdots & [p_{n-N+1}(x), x^{N-1}]_h \\ \int p_{n-N}(x) x^m h^{k-1} d\mu_0 & [p_{n-N}(x), 1]_h & \cdot & [p_{n-N}(x), x^{N-1}]_h \end{vmatrix}.$$

But taking into account that $m + N(k-1) = n - N$, the above expression becomes

$$= \frac{1}{d_n^*} \begin{vmatrix} 0 & [p_n(x), 1]_h & \cdots & [p_n(x), x^{N-1}]_h \\ \vdots & \vdots & & \vdots \\ 0 & [p_{n-N+1}(x), 1]_h & \cdots & [p_{n-N+1}(x), x^{N-1}]_h \\ \int p_{n-N}(x) x^m h^{k-1} d\mu_0 & [p_{n-N}(x), 1]_h & \cdot & [p_{n-N}(x), x^{N-1}]_h \end{vmatrix}.$$

Thus

$$[p_n^*(x), x^m h^k]_h = (-1)^N \frac{d_{n+1}^*}{d_n^*} \int p_{n-N}(x) x^m h^{k-1} d\mu_0 = (-1)^N \frac{d_{n+1}^*}{d_n^*} \int p_{n-N}^2(x) d\mu_0.$$

Case 2. If $n < N$, then we have

$$[p_n^*(x), x^n]_h = \frac{1}{d_n^*} \begin{vmatrix} [p_n(x), x^n]_h & [p_n(x), 1]_h & \cdots & [p_n(x), x^{n-1}]_h \\ \vdots & \vdots & & \vdots \\ [p_0(x), x^n] & [p_0(x), 1]_h & \cdot & [p_0(x), x^{n-1}]_h \end{vmatrix}.$$

Thus

$$[p_n^*(x), x^n]_h = \begin{cases} -\frac{d_{n+1}^*}{d_n^*}, & \text{if} \quad n \quad is \quad odd, \\ \frac{d_{n+1}^*}{d_n^*}, & \text{if} \quad n \quad is \quad even. \end{cases}$$

As a summary, we can state

Proposition 3.9. *Let $(\cdot,\cdot)_0$ be a positive definite bilinear form. Then $[\cdot,\cdot]_h$ is a positive definite bilinear form if only if*

$$\begin{cases} (-1)^N \frac{d^*_{n+1}}{d^*_n} > 0, & \textit{for} \quad n \geq N, \\ \frac{d^*_{n+1}}{d^*_n} > 0, & \textit{for} \quad n < N, \ \textit{with n even}, \\ \frac{d^*_{n+1}}{d^*_n} < 0, & \textit{for} \quad n < N, \ \textit{with n odd}. \end{cases}$$

3.2.2 Matrix representation of the extended Geronimus transformation

Let assume that $[\cdot,\cdot]_h$ is a positive definite bilinear form. We define the following symmetric matrix

$$J^* = \left([h(x)\check{p}^*_n(x), \check{p}^*_m(x)]_h\right)_{n,m=0}^{\infty},$$

where the corresponding orthonormal polynomials $(\check{p}^*_n(x))_{n\in\mathbb{N}}$ are

$$\check{p}^*_n(x) = \frac{1}{h^*_n} p^*_n(x), \qquad (h^*_n)^2 = [p^*_n(x), p^*_n(x)]_h, \quad h^*_n > 0.$$

In the standard Geronimus transformation (i.e. $h(x) = x$) there are two important facts concerning matrix factorizations [48, 148]

i) J^* can be decomposed as $J^* = CC^\top$ with C a lower triangular matrix (Cholesky factorization).

ii) If $p_n(0) \neq 0$ for $n = 0, 1, \cdots$, then there exist U, an upper triangular matrix, and L, a lower triangular matrix, such that

$$J_{mon} = UL \qquad \text{and} \qquad J^*_{mon} = LU.$$

Is it possible to extend these two results to the generalized Geronimus transformations analyzed in the previous sections?. From (3.8) we know that the polynomials $p^*_n(x)$ can be written in terms of the monic orthogonal polynomials with respect to $(\cdot,\cdot)_0$. From this we get

$$\begin{aligned}(p^*_n, p^*_m)_0 = A^{[n]}_n \sum_{i=0}^{N} A^{[m]}_{m-i}(p_n, p_{m-i})_0 + A^{[n]}_{n-1} \sum_{i=0}^{N} A^{[m]}_{m-i}(p_{n-1}, p_{m-i})_0 + \cdots \\ + A^{[n]}_{n-j} \sum_{i=0}^{N} A^{[m]}_{m-i}(p_{n-j}, p_{m-i})_0 + \cdots + A^{[n]}_{n-N} \sum_{i=0}^{N} A^{[m]}_{m-i}(p_{n-N}, p_{m-i})_0\end{aligned}$$

$$= \begin{cases} \sum_{k=d}^{N} A^{[n+d]}_{n+d-k} A^{[n]}_{n+d-k} h^2_{n-k+d}, & \text{if} \quad m = n+d, \ 0 \leq d \leq N, \\ \sum_{k=d}^{N} A^{[n]}_{n-k} A^{[n-d]}_{n-k} h^2_{n-k}, & \text{if} \quad m = n-d, \ 0 \leq d \leq N, \end{cases} \tag{3.14}$$

where $A_k^{[k]} = 1$ and $A_m^{[k]} = 0$ if $m < 0$. Notice that the above values are zero for $|n-m| \geq N$, and, therefore, J^* is a $2N+1$ diagonal matrix.

Proposition 3.10. *Let us assume that* $(\cdot,\cdot)_0$ *and* $[\cdot,\cdot]_h$ *are positive definite bilinear forms and let us denote by* $(p_n(x))_{n\in\mathbb{N}}$ *and* $(p_n^*(x))_{n\in\mathbb{N}}$*, respectively, the corresponding sequences of monic orthogonal polynomials . Then the matrix* J^* *can be written as*

$$J^* = CC^\top,$$

where C *is a lower triangular matrix with positive diagonal entries,*

$$C = \begin{bmatrix} \frac{h_0}{h_0^*} & & & & & \\ \frac{A_0^{[1]}h_0}{h_1^*} & \frac{h_1}{h_1^*} & & & & \\ \frac{A_0^{[2]}h_0}{h_2^*} & \frac{A_1^{[2]}h_1}{h_2^*} & \frac{h_2}{h_2^*} & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ 0 & \frac{A_1^{[N+1]}h_1}{h_{N+1}^*} & \dots & \frac{A_N^{[N+1]}h_N}{h_{N+1}^*} & \frac{h_{N+1}}{h_{N+1}^*} & \\ \vdots & \vdots & & & \ddots & \ddots \end{bmatrix}.$$

Proof. According to the definition of J^*

$$J^* = \begin{bmatrix} \frac{1}{h_0^*} & 0 & & \\ & \frac{1}{h_1^*} & \ddots & \\ & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} [hp_0^*(x),p_0^*(x)]_h & [hp_0^*(x),p_1^*(x)]_h & \\ [hp_1^*(x),p_0^*(x)]_h & [hp_1^*(x),p_1^*(x)]_h & \ddots \\ & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \frac{1}{h_0^*} & 0 & & \\ & \frac{1}{h_1^*} & \ddots & \\ & & \ddots & \ddots \end{bmatrix}.$$

Taking into account the definition of $[\cdot,\cdot]_h$, we have $[hp_n^*(x),p_m^*(x)]_h = (p_n^*(x),p_m^*(x))_0$. From (3.14)

$$(p_n^*(x),p_m^*(x))_0 = \begin{bmatrix} h_0^2 & A_0^{[1]}h_0^2 & A_0^{[2]}h_0^2 & \cdots & A_0^{[N]}h_0^2 & 0 \\ A_0^{[1]}h_0^2 & \sum_{k=0}^N (A_{1-k}^{[1]})^2 h_{1-k}^2 & \sum_{k=0}^N A_{3-k}^{[3]}A_{3-k}^{[2]}h_{3-k}^2 & \cdots & \sum_{k=N-1}^N A_{N-k}^{[N]}A_{N-k}^{[2]}h_{N-k}^2 & A_1^{[N+1]}h_1^2 \\ A_0^{[2]}h_0^2 & \sum_{k=1}^N A_{2-k}^{[2]}A_{2-k}^{[1]}h_{2-k}^2 & \sum_{k=0}^N (A_{2-k}^{[2]})^2 h_{2-k}^2 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \ddots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ A_0^{[N]}h_0^2 & \sum_{k=N-1}^N A_{N-k}^{[N]}A_{N-k}^{[1]}h_{N-k}^2 & \cdots & \cdots & \sum_{k=0}^N (A_{N-k}^{[N]})^2 h_{N-k}^2 & \sum_{k=1}^N A_{N+1-k}^{[N+1]}A_{N+1-k}^{[N]}h_{N+1-k}^2 \\ 0 & A_1^{[N+1]}h_1^2 & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

From the shape of the matrix, this can be written as

$$\begin{bmatrix} h_0 & & & & & \\ A_0^{[1]}h_0 & h_1 & & & & \\ A_0^{[2]}h_0 & A_1^{[2]}h_1 & h_2 & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ 0 & A_1^{[N+1]}h_1 & \cdots & A_N^{[N+1]}h_N & h_{N+1} & \\ \vdots & \vdots & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} h_0 & A_0^{[1]}h_0 & A_0^{[2]}h_0 & \cdots & \cdots & 0 & \cdots \\ & h_1 & A_1^{[2]}h_1 & \cdots & \cdots & A_1^{[N+1]}h_1 & \cdots \\ & & h_2 & \ddots & & \vdots & \\ & & & \ddots & \ddots & \vdots & \\ & & & & \ddots & A_N^{[N+1]}h_N & \cdots \\ & & & & & h_{N+1} & \cdots \\ & & & & & & \ddots \end{bmatrix}.$$

If

$$C = \begin{bmatrix} \frac{1}{h_0^*} & 0 & & \\ 0 & \frac{1}{h_1^*} & \ddots & \\ & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} h_0 & & & & & \\ A_0^{[1]}h_0 & h_1 & & & & \\ A_0^{[2]}h_0 & A_1^{[2]}h_1 & h_2 & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ 0 & A_1^{[N+1]}h_1 & \cdots & A_N^{[N+1]}h_N & h_{N+1} & \\ \vdots & \vdots & & & \ddots & \ddots \end{bmatrix},$$

then we get the desired result. Note that

$$\begin{aligned} (h_{n+N}^*)^2 &= [p_{n+N}^*(x), p_{n+N}^*(x)]_h = [hp_n^*(x), p_{n+N}^*(x)]_h \\ &= (p_n^*(x), p_{n+N}^*(x))_0 = \sum_{k=N}^{N} A_{n+N-k}^{[n+N]} A_{n+N-k}^{[n]} h_{n+N-k}^2 \\ &= A_n^{[n+N]} h_n^2. \end{aligned}$$

Hence the diagonal entries of C can be given in terms of the coefficients $A_n^{[k]}$ as follows

$$\frac{h_{n+N}}{h_{n+N}^*} = \frac{h_{n+N}}{\sqrt{A_n^{[n+N]}}h_n}.$$

In addition, if $m < N$, then

$$\begin{aligned} (h_m^*)^2 &= [p_m^*(x), p_m^*(x)]_h = [\sum_{k=0}^{m} A_k^{[m]} p_k(x), \sum_{j=0}^{m} A_k^{[m]} p_j(x)]_h \\ &= \sum_{k=0}^{m} \sum_{j=0}^{m} A_k^{[m]} A_k^{[m]} [p_k(x), p_j(x)]_h. \end{aligned}$$

From the above relation we can see that $(h_m^*)^2$ is a combination of the free parameters given in the matrix (3.4). ■

Let L_{mon} be the matrix associated with the recurrence formula given in (3.8)

$$L_{mon} = \begin{bmatrix} 1 & & & & & & \\ A_0^{[1]} & 1 & & & & & \\ A_0^{[2]} & A_1^{[2]} & 1 & & & & \\ \vdots & \vdots & \ddots & \ddots & & & \\ A_0^{[N]} & A_1^{[N]} & \cdots & A_{N-1}^{[N]} & 1 & & \\ 0 & A_1^{[N+1]} & \cdots & \cdots & A_N^{[N+1]} & 1 & \\ \vdots & \vdots & & & & \ddots & \ddots \end{bmatrix}.$$

It is clear that the relation (3.8) reads as $p^* = L_{mon}p$, where $p^* = (p_0^*(x), p_1^*(x), \cdots)^\top$ and $p = (p_0(x), p_1(x), \cdots)^\top$. It is also straightforward to show that L_{mon} is a nonsingular matrix. On the other hand,

$$[h(x)p_n(x), p_m^*(x)]_h = (p_n(x), p_m^*(x))_0 = 0, \text{ for } m = 0, \cdots, n-1.$$

Then we can write

$$h(x)p_n(x) = \sum_{i=n}^{N+n} B_i^{[N+n]} p_i^*(x), \quad \text{where } B_n^{[N+n]} \neq 0.$$

Thus we can associate with the above relation the following nonsingular matrix

$$U_{mon} = \begin{bmatrix} B_0^{[N]} & B_1^{[N]} & \cdots & \cdots & B_{N-1}^{[N]} & 1 & & & & \\ & B_1^{[N+1]} & \cdots & \cdots & B_{N-1}^{[N+1]} & B_N^{[N+1]} & 1 & & & \\ & & \ddots & & & \ddots & \ddots & \ddots & & \\ & & & B_n^{[n+N]} & & & B_{n+N-2}^{[n+N]} & B_{n+N-1}^{[n+N]} & 1 & \\ & & & & \ddots & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Here $h(x)p = U_{mon}p^*$, where p and p^* are the vectors defined as above. Finally, we can state the following

Proposition 3.11. *If $h(x) = \sum_{m=0}^{N} b_m x^m$, then*

$$h(J_{mon}) = \sum_{m=0}^{N} b_m J_{mon}^m = U_{mon}L_{mon}, \tag{3.15}$$

as well as

$$J_{mon}^* = L_{mon}U_{mon}. \tag{3.16}$$

Proof. Taking into account
$$h(x)p = U_{mon}p^* = U_{mon}L_{mon}p,$$
from
$$x^m p = J_{mon}x^{m-1}p = \cdots = J_{mon}^m p$$
we get
$$h(x)p = \sum_{m=0}^{N} b_m x^m p = \sum_{m=0}^{N} b_m J_{mon}^m p = h(J_{mon})\, p.$$
From here, and since U_{mon} and L_{mon} are nonsingular, (3.15) holds. To prove (3.16), let us notice that
$$h(x)p^* = L_{mon}h(x)p = L_{mon}U_{mon}p^*,$$
and, since
$$h(x)p^* = J_{mon}^* p^*,$$
then (3.16) follows. Notice that we have used the fact that U_{mon} and L_{mon} are both nonsingular matrices.

■

3.3 Discrete Sobolev inner products as multiple Geronimus transformations

In this section we summarize all the previous findings together with the results of [64] and present some results as a completion of the chapter.

Consider the discrete Sobolev inner product
$$\langle f,g\rangle = \int f(x)g(x)d\mu(x) + \sum_{i=1}^{M}\sum_{j=0}^{M_i} \lambda_{i,j} f^{(j)}(x_i)g^{(j)}(x_i),$$
where f, g are polynomials and $\lambda_{i,j}$ are real numbers. We also assume that the inner product $\langle\cdot,\cdot\rangle$ is symmetric i.e. $\langle f,g\rangle = \langle g,f\rangle$. Then the following result holds.

Proposition 3.12. *The discrete Sobolev inner product $\langle\cdot,\cdot\rangle$ is a multiple Geronimus transformation of a bilinear form generated by the measure $d\mu_0(x) = h(x)d\mu(x)$, where*
$$h(x) = \prod_{i=1}^{M}(x-x_i)^{M_i+1},$$
that is
$$\langle f,g\rangle \equiv [f,g]_h.$$

Proof. This statement is a straightforward consequence of Proposition 3.2 and Example 2.11. ∎

This result, together with Proposition 3.11, gives us an understanding of the structure of the band matrices associated with the recurrence relations generated by Sobolev orthogonal polynomials.

Proposition 3.13. *Let us consider a discrete Sobolev inner product $\langle \cdot, \cdot \rangle$. Then the band matrix J^*_{mon} generated by the recurrence relations for the corresponding orthogonal polynomials can be obtained as follows*

$$h(J_{mon}) = U_{mon} L_{mon} \mapsto J^*_{mon} = L_{mon} U_{mon}, \tag{3.17}$$

where J_{mon} is the monic Jacobi matrix associated with $d\mu_0$.

Let $p(x) = \sum_{j=0}^{n} \sum_{k=0}^{N-1} a_{k,j} x^k h^j(x)$ be a polynomial of degree $nN+m$, $0 \leq m < N$, where we assume $a_{k,n} = 0$ if $k > m$. For $0 \leq k < N-1$, let us consider the linear operator $R_{k,h}(\cdot)$ introduced in Chapter 3 (see Definition 2.9).

Using the previous notation, Proposition 3.13 can be read as a result for matrix orthogonal polynomials due to [64]. Indeed, the matrix $h(J_{mon})$ generates matrix polynomials

$$P_n(x) = \begin{bmatrix} R_{0,h}(p_{nN})(x) & \dots & R_{N-1,h}(p_{nN})(x) \\ R_{1,h}(p_{nN+1})(x) & \dots & R_{N-1,h}(p_{nN+1})(x) \\ \vdots & & \vdots \\ R_{N-1,h}(p_{nN+N-1})(x) & \dots & R_{N-1,h}(p_{nN+N-1})(x) \end{bmatrix}$$

which are orthogonal with respect to the matrix of measures $dM_0(h^{-1})$, where

$$dM_0(x) = \begin{bmatrix} d\mu_0(x) & xd\mu_0(x) & \dots & x^{N-1} d\mu_0(x) \\ xd\mu_0(x) & x^2 d\mu_0(x) & \dots & x^N d\mu_0(x) \\ x^2 d\mu_0(x) & x^3 d\mu_0(x) & \dots & x^{N+1} d\mu_0(x) \\ \vdots & \vdots & & \vdots \\ x^{N-1} d\mu_0(x) & x^N d\mu_0(x) & \dots & x^{2N-2} d\mu_0(x) \end{bmatrix}$$

and $(p_n(x))_{n\in\mathbb{N}}$ is the sequence of polynomials orthogonal with respect to $d\mu_0$. At the same time, the matrix J^*_{mon} corresponds to Sobolev orthogonal polynomials which, in turn, can be considered as matrix orthogonal polynomials with respect to the matrix of measures (see Example 2.11)

$$dM(h^{-1}) + L\delta_0, \quad \text{where} \quad dM(x) = \begin{bmatrix} d\mu(x) & xd\mu(x) & \dots & x^{N-1} d\mu(x) \\ xd\mu(x) & x^2 d\mu(x) & \dots & x^N d\mu(x) \\ x^2 d\mu(x) & x^3 d\mu(x) & \dots & x^{N+1} d\mu(x) \\ \vdots & \vdots & & \vdots \\ x^{N-1} d\mu(x) & x^N d\mu(x) & \dots & x^{2N-2} d\mu(x) \end{bmatrix} \tag{3.18}$$

and L is the matrix

$$\sum_{i=1}^{M} \sum_{j=0}^{M_i} \lambda_{i,j} \mathbf{L}(i,j),$$

where $\mathbf{L}(i,j)$ is the $N\times N$ matrix

$$\mathbf{L}(i,j)=\begin{bmatrix}0\\ \vdots\\ 0\\ j!\\ \vdots\\ \frac{k!}{(k-j)!}x_i^{k-j}\\ \vdots\\ \frac{(N-1)!}{(N-1-j)!}x_i^{N-1-j}\end{bmatrix}\begin{bmatrix}0\dots 0 \;\; j!\dots \frac{k!}{(k-j)!}x_i^{k-j}\dots \frac{(N-1)!}{(N-1-j)!}x_i^{N-1-j}\end{bmatrix}.$$

In other words, according to (3.17), the matrix of measures (3.18) is actually a simple matrix Geronimus transformation of the matrix of measures dM_0. Thus, a multiple Geronimus transformation yields a simple Geronimus transformation for matrices of measures.

3.14 Example . *If in Example 3.7 we take $N=2$ and $d\mu_0=x^{\alpha+2}e^{-x}dx$, then the bilinear form $B(\cdot,\cdot)$, satisfying $B(x^2f,g)=B(f,gx^2)=\int_0^\infty f(x)g(x)x^{\alpha+2}e^{-x}dx$ has the following explicit form*

$$B(f,g)=\int_0^\infty f(x)g(x)x^\alpha e^{-x}dx+C_0f(0)g(0)+C_1f^{(1)}(0)g^{(1)}(0).$$

Using the results of Example 2.12, we get that the sequence of matrix polynomials $(\mathcal{L}_n(x))_{n\in\mathbb{N}}$ with

$$\mathcal{L}_n(x)=\begin{bmatrix}R_{h,0}(\tilde{L}^\alpha_{2n,2})(x) & R_{h,1}(\tilde{L}^\alpha_{2n,2})(x)\\ R_{h,0}(\tilde{L}^\alpha_{2n+1,2})(x) & R_{h,1}(\tilde{L}^\alpha_{2n+1,2})(x)\end{bmatrix}$$

is orthonormal with respect to the matrix inner product

$$\langle P(x),Q(x)\rangle=\int_0^\infty P(x)\begin{bmatrix}1 & x^{1/2}\\ x^{1/2} & x\end{bmatrix}Q(x)^\dagger e^{-x^{1/2}}x^{\alpha/2}dx+P(0)\begin{bmatrix}C_0 & 0\\ 0 & C_1\end{bmatrix}Q(0)^\dagger.$$

Thus, a multiple Geronimus transformation of a Laguerre measure yields to a simple Matrix Geronimus transformation of the "Laguerre" matrix of measures $\begin{bmatrix}1 & x^{1/2}\\ x^{1/2} & x\end{bmatrix}e^{-x^{1/2}}x^{\alpha/2}dx$.

3.4 An extension of the Geronimus transformation for orthogonal matrix polynomial on the real line

Subsection 3.3 illustrates a close relation between the multiple Geronimus transformation and a simple matrix Geronimus transformation. This allows to ask how a Matrix Geronimus transformation acts on a positive definite matrix of measures $\vartheta=(\mu_{i,j})_{i,j=0}^{p-1}$, i.e., given a monic matrix polynomial $W(x)$ and a sesquilinear form $\langle\cdot,\cdot\rangle_W$ defined from a matrix of measures ϑ as $\langle f(x)W(x),g(x)W(x)\rangle_W=\int fd\vartheta g^\dagger$, what is the relation between the sequences of orthogonal polynomials with respect to $\langle\cdot,\cdot\rangle_W$ and $d\vartheta$?. Notice that this problem is more difficult than in the scalar case because we do not assume that $\langle f(x)W(x),g(x)\rangle_W=\langle f(x),g(x)W(x)\rangle_W$, besides of the fact, that the product of matrix polynomials is noncommutative. Observe that we can also study the sesquilinear forms defined as $\langle f(x),g(x)W(x)\rangle_W=\int fd\vartheta g^\dagger$ and $\langle f(x)W(x),g(x)\rangle_W=\int fd\vartheta g^\dagger$. They will be studied in Chapter 5 but from the point of view of distributional sesquilinear forms (see Definition 1.48).

3.4.1 Geronimus transformation for orthogonal matrix polynomials on the real line

Let $\vartheta = (\mu_{i,j})_{i,j=0}^{p-1}$ be a positive definite matrix of measures supported on $\mathfrak{I} \subset \mathbb{R}$, and let us introduce the following sesquilinear form

$$\langle P(x), Q(x)\rangle_L = \int_{\mathfrak{I}} P(x) d\vartheta Q^{\dagger}(x), \quad P(x), Q(x) \in \mathbb{C}^{p\times p}[x]. \tag{3.19}$$

Let M be the matrix of moments associated with $\langle\cdot,\cdot\rangle_L$. Since ϑ is a positive definite matrix of measures, then M is a positive definite matrix. This means that M has a Cholesky block factorization $M = S^{-1}HS^{-\dagger}$, where S a lower triangular block matrix with I_p in its main diagonal and H is a block diagonal matrix. The sequence of matrix monic orthogonal polynomials with respect to (3.19) is given by $P = S\chi(x)$ where $P = [P_0(x)^{\dagger}, P_1(x)^{\dagger}, \cdots]^{\dagger}$.

Let $W(x)$ be a matrix monic polynomial of degree N with Np zeros outside the interior of the convex hull of $\mathfrak{I}$, the support of the matrix of measures $d\vartheta$. Let $\mathcal{B}_W$ be the set $\mathcal{B}_W := \{x^iW^m : i = 0,\dots,N-1, m > 0\}$. Since $\deg(x^iW^m) = i+Nm$, if we denote by $r_{Nm+l}(x) = x^lW^m(x)$, with $m \geq 0$ and $0 \leq l \leq N-1$, then $(r_n(x))_{n\in\mathbb{N}}$ is another ordered basis of monic polynomials of the left module $\mathbb{C}^{p\times p}[x]$, $deg(r_n(x)) = n$. Thus, if $M_{\mathcal{B}}$ is the matrix of moments associated with this new basis, i.e. $(M_{\mathcal{B}})_{i,j} =: \mu_{i,j} = \langle r_i(x), r_j(x)\rangle_L$ and $(M_{\mathcal{B}})_n$ is its $(n-1)$-th block truncation, then from the uniqueness of $(P_n(x))_{n\in\mathbb{N}}$ it is easy to see that

$$P_n(x) = r_n(x) - \begin{bmatrix}\mu_{n,0} & \mu_{n,1} & \cdots & \mu_{n,n-1}\end{bmatrix} (M_{\mathcal{B}})_n^{-1} \begin{bmatrix} r_0(x) \\ r_1(x) \\ \vdots \\ r_{n-1}(x)\end{bmatrix}. \tag{3.20}$$

In the same way, the kernel matrix polynomial has a representation in terms of the moments associated with the basis $(r_n(x))_{n\in\mathbb{N}}$ as follows

$$K_n(x,y) = \begin{bmatrix} r_0^{\dagger}(y) & \dots & r_n^{\dagger}(y)\end{bmatrix} (M_{\mathcal{B}})_{n+1}^{-1} \begin{bmatrix} r_0(x) \\ \vdots \\ r_n(x)\end{bmatrix}. \tag{3.21}$$

Now, we define a sesquilinear form $\langle f, g\rangle_W$ on $\mathbb{C}^{p\times p}[x]$ such that

$$\langle R(x)W(x), Q(x)W(x)\rangle_W = \int_{\mathfrak{I}} R(x) d\vartheta Q(x)^{\dagger}. \tag{3.22}$$

Notice that $\langle\cdot,\cdot\rangle_W$ is not completely defined by (3.22). Indeed, if $\check{\mu}_{k,j}$ are the moments associated with $\langle\cdot,\cdot\rangle_W$ with respect to the basis $\mathcal{B}_W$, i.e. $\check{\mu}_{Nm+k,Nm'+k'} = \left\langle x^kW^m(x), x^{k'}W^{m'}(x)\right\rangle_W$, then for $0 \leq k,k' \leq N-1$, the moments $\check{\mu}_{k,Nm'+k'}$ and $\check{\mu}_{Nm+k,k'}$ (that is, the first N rows and columns on the matrix of moments) can be chosen arbitrarily. However, we require that $\check{\mu}_{Nm+k,k'}$=$\check{\mu}^{\dagger}_{k',Nm+k}$ in order that $\langle\cdot,\cdot\rangle_W$ will be hermitian. If $\langle\cdot,\cdot\rangle_W$ is a quasi-definite sesquilinear form and we denote by $M_{\mathcal{B}}$ and $\check{M}_{\mathcal{B}}$ the block moment matrices associated with $d\vartheta$ and $\langle\cdot,\cdot\rangle_W$, respectively, using this basis, then $\check{M}_{\mathcal{B}}$ and $M_{\mathcal{B}}$ are related as follows

$$M_{\mathcal{B}} = \begin{bmatrix} \mu_{0,0} & \mu_{0,1} & \mu_{0,2} & \cdots \\ \mu_{1,0} & \mu_{1,1} & \mu_{1,2} & \cdots \\ \mu_{2,0} & \mu_{2,1} & \mu_{2,2} & \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \check{M}_{\mathcal{B}} = \left[\begin{array}{ccc|c} \check{\mu}_{0,0} & \cdots & \check{\mu}_{0,N-1} & \cdots \\ \vdots & \ddots & \vdots & \\ \check{\mu}_{N-1,0} & \cdots & \check{\mu}_{N-1,N-1} & \cdots \\ \hline \vdots & & \vdots & M_{\mathcal{B}} \end{array}\right].$$

On the other hand, the matrix of moments $\check{M}$ associated with $\langle\cdot,\cdot\rangle_W$, computed in terms of the canonical basis $(x^n I_p)_{n\in\mathbb{N}}$ has a block Cholesky factorization $\check{M} = \check{S}^{-1}\check{H}\check{S}^{-\dagger}$ with $\check{S}$ a lower triangular block matrix with I_p in its main diagonal and $\check{H}$ a block diagonal matrix. The block matrices M and $\check{M}$ are related as follows

$$M = \langle\chi(x),\chi(x)\rangle_L = \langle\chi(x)W(x),\chi(x)W(x)\rangle_W = W(\Lambda)\,\langle\chi(x),\chi(x)\rangle_W\,W(\Lambda)^\dagger = W(\Lambda)\check{M}W(\Lambda)^\dagger.$$

For the existence of the sequence of monic matrix orthogonal polynomials $(\check{P}_n(x))_{n\in N}$ with respect to $\langle\cdot,\cdot\rangle_W$ we need that the matrices $(\check{M}_{\mathcal{B}})_n$ must be nonsingular for every $n\in N$. Indeed, if we consider the quasi-determinant

$$|(\check{M}_{\mathcal{B}})_{N+n}| =: \Theta_*\begin{bmatrix} \boxed{(\check{M}_{\mathcal{B}})_N} & E_n \\ E_n^\dagger & (M_{\mathcal{B}})_n \end{bmatrix} \quad \text{with} \quad E_n = \begin{bmatrix} \check{\mu}_{0,N} & \cdots & \cdots & \check{\mu}_{0,N+n-1} \\ \vdots & & & \vdots \\ \check{\mu}_{N-1,N} & \cdots & \cdots & \check{\mu}_{N-1,N+n-1} \end{bmatrix},$$

then using the determinant formula $\det((\check{M}_{\mathcal{B}})_{N+n}) = \det((M_{\mathcal{B}})_N)\det(|(\check{M}_{\mathcal{B}})_{N+n}|)$ we conclude that the sequence of polynomials $(\check{P}_n(x))_{n\in N}$ will exist if the matrices $(\check{M}_{\mathcal{B}})_k$, $k=1,\dots,N$, and $|(\check{M}_{\mathcal{B}})_{N+n}|$, $n\in\mathbb{N}$, are nonsingular. Observe that for $n = N(l-1)+s$, with $s=0,\dots,N-1$, and $l>1$ we have

$$|(\check{M}_{\mathcal{B}})_{N+n}| = \Theta_*\left[\begin{array}{c|c|c} \boxed{(\check{M}_{\mathcal{B}})_N} & A_n & B_n \\ \hline A_n^\dagger & (M_{\mathcal{B}})_{n-N} & C_n \\ \hline B_n^\dagger & C_n^\dagger & D_n \end{array}\right], \tag{3.23}$$

where

$$A_n = \begin{bmatrix} \check{\mu}_{0,N} & \cdots & \check{\mu}_{0,n-1} \\ \vdots & & \vdots \\ \check{\mu}_{N-1,N} & \cdots & \check{\mu}_{N-1,n-1} \end{bmatrix}, \quad B_n = \begin{bmatrix} \check{\mu}_{0,n} & \cdots & \check{\mu}_{0,n+N-1} \\ \vdots & & \vdots \\ \check{\mu}_{N-1,n} & \cdots & \check{\mu}_{N-1,n+N-1} \end{bmatrix},$$

$$C_n = \begin{bmatrix} \mu_{0,n-N} & \cdots & \mu_{0,n-1} \\ \vdots & & \vdots \\ \mu_{n-N-1,n-N} & \cdots & \check{\mu}_{n-N-1,n-1} \end{bmatrix}, \quad D_n = \begin{bmatrix} \mu_{n-N,n-N} & \cdots & \mu_{n-N,n-1} \\ \vdots & & \vdots \\ \mu_{n-1,n-N} & \cdots & \check{\mu}_{n-1,n-1} \end{bmatrix}.$$

With this in mind, we get the following proposition.

Proposition 3.15. *For $\ell > 1$ and $0 \le s \le N-1$,*

$$|(\check{M}_{\mathcal{B}})_{N\ell+s}| = (\check{M}_{\mathcal{B}})_N + E_s (M_{\mathcal{B}})_s^{-1} E_s^{\dagger} - \tag{3.24}$$
$$\sum_{j=1}^{\ell-1} \Theta_* \begin{bmatrix} \boxed{B_{Nj+s}} & A_{Nj+s} \\ C_{Nj+s} & (M_{\mathcal{B}})_{N(j-1)+s} \end{bmatrix} \Theta_* \begin{bmatrix} \boxed{D_{Nj+s}} & C^{\dagger}_{Nj+s} \\ C_{Nj+s} & (M_{\mathcal{B}})_{N(j-1)+s} \end{bmatrix}^{-1} \Theta_* \begin{bmatrix} \boxed{B_{Nj+s}} & C^{\dagger}_{Nj+s} \\ A^{\dagger}_{Nj+s} & (M_{\mathcal{B}})_{N(j-1)+s} \end{bmatrix}.$$

Moreover,

$$\Theta_* \begin{bmatrix} \boxed{B_n} & C_n^{\dagger} \\ A_n^{\dagger} & (M_{\mathcal{B}})_{n-N} \end{bmatrix} = \Theta_* \begin{bmatrix} \boxed{B_n} & A_n \\ C_n & (M_{\mathcal{B}})_{n-N} \end{bmatrix}^{\dagger} =: \begin{bmatrix} d_{n,0} & \cdots & d_{n,n-1} \\ \vdots & & \vdots \\ d_{n+N-1,0} & \cdots & d_{n+N-1,N-1} \end{bmatrix},$$

where

$$d_{m+N,k} = \left\langle P_m(x)W(x) + \left[\left\langle r_m(y), K^{\dagger}_{m-1}(x,y) - K^{\dagger}_{n-N-1}(x,y)\right\rangle_L\right] W(x), r_k(x)\right\rangle_W,$$

for $n-N \le m \le n-1$, $0 \le k \le N-1$, and

$$\Theta_* \begin{bmatrix} \boxed{D_n} & C_n^{\dagger} \\ C_n & (M_{\mathcal{B}})_{n-N} \end{bmatrix} =: \begin{bmatrix} h_{n-N,n-N} & \cdots & h_{n-N,n-1} \\ \vdots & & \vdots \\ h_{n-1,n-N} & \cdots & h_{n-1,n-1} \end{bmatrix},$$

where

$$h_{m,k} = \left\langle P_m(x) + \left[\left\langle r_m(y), K^{\dagger}_{m-1}(x,y) - K^{\dagger}_{n-N-1}(x,y)\right\rangle_L\right], r_k(x)\right\rangle_L,$$

for $n-N \le m,k \le n-1$.

Proof. Let $n \ge N$. From (3.23)

$$|(\check{M}_{\mathcal{B}})_{N+n}| = \Theta_* \left[\begin{array}{c|c|c} \boxed{(\check{M}_{\mathcal{B}})_N} & B_n & A_n \\ \hline B_n^{\dagger} & D_n & C_n^{\dagger} \\ \hline A_n^{\dagger} & C_n & (M_{\mathcal{B}})_{n-N} \end{array}\right].$$

Since B_n and D_n are square matrices, then using Sylvester's theorem (see (1.6)), we get

$$|(\check{M}_{\mathcal{B}})_{N+n}| =$$
$$|(\check{M}_{\mathcal{B}})_n| - \Theta_* \begin{bmatrix} \boxed{B_n} & A_n \\ C_n & (M_{\mathcal{B}})_{n-N} \end{bmatrix} \Theta_* \begin{bmatrix} \boxed{D_n} & C_n^{\dagger} \\ C_n & (M_{\mathcal{B}})_{n-N} \end{bmatrix}^{-1} \Theta_* \begin{bmatrix} \boxed{B_n} & C_n^{\dagger} \\ A_n^{\dagger} & (M_{\mathcal{B}})_{n-N} \end{bmatrix}.$$

Thus, (3.24) follows in a recursive way. On the other hand, for $n-N \leq m \leq n-1$ and $0 \leq k \leq n-1$,

$$d_{m+N,k} = \check{\mu}_{m+N,k} - (\mu_{m,0}, \cdots, \mu_{m,n-N-1}) \left[(M_{\mathcal{B}})_{n-N}\right]^{-1} \begin{bmatrix} \check{\mu}_{N,k} \\ \vdots \\ \check{\mu}_{n-1,k} \end{bmatrix}$$

and using (3.20) and (3.21) we get

$$\check{\mu}_{m+N,k} = \left\langle \left[P_m(x) + \left\langle r_m(y), K^{\dagger}_{m-1}(x,y) \right\rangle_L\right] W(x), r_k(x) \right\rangle_W,$$

$$(\mu_{m,0}, \cdots, \mu_{m,n-N-1}) \left[(M_{\mathcal{B}})_{n-N}\right]^{-1} \begin{bmatrix} \check{\mu}_{N,k} \\ \vdots \\ \check{\mu}_{n-1,k} \end{bmatrix} = \left\langle \left[\left\langle r_m(y), K^{\dagger}_{m-1}(x,y) - K^{\dagger}_{n-N-1}(x,y) \right\rangle_L\right] W(x), r_k(x) \right\rangle_W.$$

In the same way, for $n-N \leq m,k \leq n-1$,

$$h_{m,k} = \mu_{m,k} - (\mu_{m,0}, \cdots, \mu_{m,n-N-1}) \left[(M_{\mathcal{B}})_{n-N}\right]^{-1} \begin{bmatrix} \mu_{0,k} \\ \vdots \\ \mu_{n-N-1,k} \end{bmatrix},$$

and thus we get the result. ■

Recall that since $\mathcal{B}_W$ is a basis of the left module $\mathbb{C}^{p\times p}[x]$, every matrix polynomial $f(x)$ of degree $n = sN + j$ can be written as

$$f(x) = \sum_{l=0}^{N-1} \sum_{m=0}^{s} a_{l,m} x^l W^m(x), \tag{3.25}$$

where $a_{l,m} = 0_p$ if $m = s$ and $l > j$. Let x_i be a zero of $W(x)$ with multiplicity α_i, $k = 1, \cdots, q$. If $\{r^{(i)}_{1,0} \ldots, r^{(i)}_{1,s_i}\}$ is a basis of $\operatorname{Ker}(W(x_i))$ with dimension s_i, then there exists a canonical Jordan chain

$$r^{(i)}_{j,0}, r^{(i)}_{j,1}, \cdots, r^{(i)}_{j,\kappa^{(i)}_j-1}, \quad j = 1, \ldots, s_i,$$

with $\sum_{j=1}^{s_i} \kappa^{(i)}_j = \alpha_i$ and the corresponding right root polynomials $r^{(i)}_j(x)$. With this in mind, and from Definition 1.35, Proposition 1.33, and the representation (3.25), we obtain

$$\mathcal{J}_f(x_i) = a_{0,0}\mathcal{J}_{I_p}(x_i) + \cdots + a_{N-1,0}\mathcal{J}_{I_p x^{N-1}}(x_i) = [a_{0,0} \cdots a_{N-1,0}] \begin{bmatrix} Q^{(i)}_0 \\ \vdots \\ Q^{(i)}_{N-1} \end{bmatrix}.$$

As a consequence,

$$\mathcal{J}_f = [a_{0,0} \cdots a_{N-1,0}] Q,$$

and since Q is a nonsingular matrix (see Lemma 1.40), then

$$\mathcal{J}_f Q^{-1} = [a_{0,0} \cdots a_{N-1,0}].$$

On the other hand, let $f^{[1]}(x)$ be the matrix polynomial defined by

$$f^{[1]}(x) = \left(f(x) - \sum_{l=0}^{N-1} a_{l,0} x^l \right) W^{-1}(x).$$

In a similar way, we define recursively the following sequence of matrix polynomials $(f^{[k]}(x))_{k=2}^{s}$

$$f^{[k]}(x) = \left(f^{[k-1]}(x) - \sum_{l=0}^{N-1} a_{l,k-1} x^l \right) W^{-1}(x) = \sum_{m=k}^{s} \sum_{l=0}^{N-1} a_{l,m} x^l W^{m-k}(x).$$

Proceeding as above, for the sequence of matrix polynomials $(f^{[k]}(x))_{k=1}^{s}$ we get

$$\mathcal{J}_{f^{[k]}} Q^{-1} = [a_{0,k} \cdots a_{N-1,k}]. \tag{3.26}$$

We are now ready to obtain an explicit representation for $\langle \cdot, \cdot \rangle_W$.

Proposition 3.16. *Let $d\check{\vartheta}$ be a positive definite matrix of measures such that $W(x) d\check{\vartheta} W(x)^\dagger = d\vartheta$. Let $f = \sum_{l=0}^{N-1} \sum_{m=0}^{s} a_{l,m} x^l W^m(x)$ and $g = \sum_{l'=0}^{N-1} \sum_{m'=0}^{s'} a_{l',m'} x^{l'} W^{m'}(x)$ be arbitrary matrix polynomials. Then $\langle \cdot, \cdot \rangle_W$ can be represented as follows*

$$\begin{aligned}
\langle f, g \rangle_W &= \int f d\check{\vartheta} g^\dagger + \sum_{m=1}^{r} \mathcal{J}_{f^{[m]}} Q^{-1} \begin{pmatrix} \check{\Omega}_{Nm,0} & \cdots & \check{\Omega}_{Nm,N-1} \\ \vdots & & \\ \check{\Omega}_{N-1+Nm,0} & \cdots & \check{\Omega}_{N-1+Nm,N-1} \end{pmatrix} Q^{-\dagger} \mathcal{J}_g^\dagger \\
&+ \sum_{m=1}^{r} \mathcal{J}_f Q^{-1} \begin{pmatrix} \check{\Omega}_{0,Nm} & \cdots & \check{\Omega}_{0,N-1+Nm} \\ \vdots & & \\ \check{\Omega}_{N-1,Nm} & \cdots & \check{\Omega}_{N-1,N-1+Nm} \end{pmatrix} Q^{-\dagger} \mathcal{J}_{g^{[m]}}^\dagger + \mathcal{J}_f Q^{-1} \begin{pmatrix} \check{\Omega}_{0,0} & \cdots & \check{\Omega}_{0,N-1} \\ \vdots & & \\ \check{\Omega}_{N-1,0} & \cdots & \check{\Omega}_{N-1,N-1} \end{pmatrix} Q^{-\dagger} \mathcal{J}_g^\dagger,
\end{aligned}$$

where $r = \max\{s, s'\}$, and

$$\check{\Omega}_{i+Nj,i'+Nj'} = \left\langle x^i W^j(x), x^{i'} W^{j'}(x) \right\rangle_W - \int x^i W^j(x) d\check{\vartheta} \left(x^{i'} W^{j'}(x) \right)^\dagger, \tag{3.27}$$

i.e., the difference between the moments associated with the bilinear form $\langle \cdot, \cdot \rangle_W$ and the moments associated with $d\check{\vartheta}$.

Remark 3.17. *For $m = 0, \ldots, r$, the matrices*

$$\begin{pmatrix} \check{\Omega}_{Nm,0} & \cdots & \check{\Omega}_{Nm,N-1} \\ \vdots & & \\ \check{\Omega}_{N-1+Nm,0} & \cdots & \check{\Omega}_{N-1+Nm,N-1} \end{pmatrix}, \quad \begin{pmatrix} \check{\Omega}_{0,Nm} & \cdots & \check{\Omega}_{0,Nm+N-1} \\ \vdots & & \\ \check{\Omega}_{N-1,Nm} & \cdots & \check{\Omega}_{N-1,Nm+N-1} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \check{\Omega}_{0,0} & \cdots & \check{\Omega}_{0,N-1} \\ \vdots & & \\ \check{\Omega}_{N-1,0} & \cdots & \check{\Omega}_{N-1,N-1} \end{pmatrix}$$

depend on the the moments $\check{\mu}_{k,Nm'+k'}$ and $\check{\mu}_{Nm+k,k'}$, which can be arbitrarily chosen for $0 \leq k, k' \leq N-1$.

Proof. Let us write

$$\langle f,g\rangle_W=\left\langle\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}x^lW^m(x),\sum_{l'=0}^{N-1}\sum_{m'=1}^{s'}b_{l',m'}x^{l'}W^{m'}(x)\right\rangle_W+\left\langle\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}x^lW^m(x),\sum_{l'=0}^{N-1}b_{l',0}x^{l'}\right\rangle_W$$
$$+\left\langle\sum_{l=0}^{N-1}a_{l,0}x^l,\sum_{l'=0}^{N-1}\sum_{m'=1}^{s'}a_{l',m'}x^{l'}W^{m'}(x)\right\rangle_W+\left\langle\sum_{l=0}^{N-1}a_{l,0}x^lW^m(x),\sum_{l'=0}^{N-1}b_{l',0}x^{l'}\right\rangle_W$$
$$=\int fd\check{\vartheta}g^\dagger+\left\langle\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}x^lW^m(x),\sum_{l'=0}^{N-1}b_{l',0}x^{l'}\right\rangle_W-\int\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}x^lW^m(x)d\check{\vartheta}\left(\sum_{l'=0}^{N-1}b_{l',0}x^{l'}\right)^\dagger$$
$$+\left\langle\sum_{l=0}^{N-1}a_{l,0}x^l,\sum_{l'=0}^{N-1}\sum_{m'=1}^{s'}b_{l',m'}x^{l'}W^{m'}(x)\right\rangle_W-\int\sum_{l=0}^{N-1}a_{l,0}x^ld\check{\vartheta}\left(\sum_{l'=0}^{N-1}\sum_{m'=1}^{s'}b_{l',m'}x^{l'}W^{m'}(x)\right)^\dagger$$
$$+\left\langle\sum_{l=0}^{N-1}a_{l,0}x^lW^m(x),\sum_{l'=0}^{N-1}b_{l',0}x^{l'}\right\rangle_W-\int\sum_{l=0}^{N-1}a_{l,0}x^ld\check{\vartheta}\left(\sum_{l'=0}^{N-1}b_{l',0}x^{l'}\right)^\dagger.$$

Defining $\check{\Omega}_{i+Nj,i'+Nj'}$ as in (3.27), we get

$$\sum_{l'=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}\check{\Omega}_{l+Nm,l'}b^\dagger_{l',0}=$$
$$=\begin{bmatrix}\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}\check{\Omega}_{l+Nm,0}&\cdots&\sum_{l=0}^{N-1}\sum_{m=1}^{s}a_{l,m}\check{\Omega}_{l+Nm,N-1}\end{bmatrix}\begin{bmatrix}b^\dagger_{0,0}\\\vdots\\b^\dagger_{N-1,0}\end{bmatrix}$$
$$=\sum_{m=1}^{s}\left(\begin{bmatrix}a_{0,m}&\cdots&a_{N-1,m}\end{bmatrix}\begin{bmatrix}\check{\Omega}_{Nm,0}\\\vdots\\\check{\Omega}_{N-1+Nm,0}\end{bmatrix}\cdots\begin{bmatrix}a_{0,m}&\cdots&a_{N-1,m}\end{bmatrix}\begin{bmatrix}\check{\Omega}_{Nm,N-1}\\\vdots\\\check{\Omega}_{N-1+Nm,N-1}\end{bmatrix}\right)\begin{bmatrix}b^\dagger_{0,0}\\\vdots\\b^\dagger_{N-1,0}\end{bmatrix}$$
$$=\sum_{m=1}^{s}\begin{bmatrix}a_{0,m}&\cdots&a_{N-1,m}\end{bmatrix}\begin{bmatrix}\check{\Omega}_{Nm,0}&\cdots&\check{\Omega}_{Nm,N-1}\\\vdots&&\\\check{\Omega}_{N-1+Nm,0}&\cdots&\check{\Omega}_{N-1+Nm,N-1}\end{bmatrix}\begin{bmatrix}b^\dagger_{0,0}\\\vdots\\b^\dagger_{N-1,0}\end{bmatrix}.$$

Similarly,

$$\sum_{l'=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m'=1}^{s}a_{l,0}\check{\Omega}_{l,Nm'+l'}b^\dagger_{l',m'}$$
$$=\sum_{m'=1}^{s'}\begin{bmatrix}a_{0,0}&\cdots&a_{N-1,0}\end{bmatrix}\begin{bmatrix}\check{\Omega}_{0,Nm'}&\cdots&\check{\Omega}_{0,Nm'+N-1}\\\vdots&&\\\check{\Omega}_{N-1,Nm'}&\cdots&\check{\Omega}_{N-1,Nm'+N-1}\end{bmatrix}\begin{bmatrix}b^\dagger_{0,m'}\\\vdots\\b^\dagger_{N-1,m'}\end{bmatrix},$$

and taking into account the previous equations and (3.26), the result follows. The limit on the sums can be taken as $r = \max\{s, s'\}$ since the additional terms vanish. ■

Corollary 3.18. *If $W(x) = (xI_p - A)$, then*

$$\langle f, g\rangle_W = \int f d\check{\vartheta} g^\dagger + \sum_{i=1}^{s} f^{(i)}(A)\frac{1}{i!}\check{\Omega}_{i,0}[g(A)]^\dagger + \sum_{i=1}^{s} f(A)\frac{1}{i!}\check{\Omega}_{0,i}[g^{(i)}(A)]^\dagger + f(A)\check{\Omega}_{0,0}[g(A)]^\dagger.$$

Proof. Given a polynomial $f(x)$ of degree s, it can be written as $\sum_{m=0}^{s} a_m(xI_p - A)^m$, where $a_m \in \mathbb{C}^{p\times p}$, $0 \le m \le s$. Notice that since $W(x) = (xI_p - A)$, then from (3.26) we get $a_m = \mathcal{J}_{f^{[m]}} Q^{-1}$. On the other hand, it is clear that $a_m = \frac{1}{m!} f^{(m)}(A)$. The above yields the result. Observe the connection between the Jordan chain and the evaluation of a polynomial at a matrix. ■

3.5 Connection formulas

Let $(\check{P}_n(x))_{n\in\mathbb{N}}$ be the sequence of monic orthogonal matrix polynomials with respect to $\langle\cdot,\cdot\rangle_W$. Using the basis $\mathcal{B}_W$ we can write $\check{P}_{n+N}(x)$ as

$$\check{P}_{n+N}(x) = \sum_{m\ge 0}\sum_{l=0}^{N-1} a_{l,m}^{[n+N]} x^l W^m(x).$$

Let $M_{\mathcal{B}}$ and $\check{M}_{\mathcal{B}}$ be the moment matrices with respect $\langle\cdot,\cdot\rangle_L$ and $\langle\cdot,\cdot\rangle_W$, respectively, computed in terms of the basis $\mathcal{B}_W$. Denoting $r_{Nm+l}(x) =: x^l W^m(x)$, $0 \le l \le N-1$, and using the same idea as in (3.20), we have

$$\check{P}_{n+N}(x) = r_{n+N}(x) - \begin{bmatrix} \check{\mu}_{n+N,0} & \check{\mu}_{n+N,1} & \cdots & \check{\mu}_{n+N,N+n-1}\end{bmatrix} (\check{M}_{\mathcal{B}})_{n+N}^{-1} \begin{bmatrix} r_0(x) \\ r_1(x) \\ \vdots \\ r_{n+N-1}(x)\end{bmatrix}.$$

Furthermore, we have

$$(\check{M}_{\mathcal{B}})_{n+N} = \begin{bmatrix} (\check{M}_{\mathcal{B}})_N & \tilde{E}_n \\ \tilde{E}_n^\dagger & (M_{\mathcal{B}})_n \end{bmatrix}, \quad \text{where} \quad \tilde{E}_n = \begin{bmatrix} \check{\mu}_{0,N} & \cdots & \cdots & \check{\mu}_{0,N+n-1} \\ \vdots & & & \vdots \\ \check{\mu}_{N-1,N} & \cdots & \cdots & \check{\mu}_{N-1,N+n-1}\end{bmatrix}.$$

With this in mind and using the inverse formula for 2×2 block matrices obtained from the Schur complement (see [83]), we obtain

$$\mathcal{J}_{\check{P}_{n+N}} = \left(\begin{bmatrix} \check{\mu}_{n+N,0} & \cdots & \check{\mu}_{n+N,N-1}\end{bmatrix} - \begin{bmatrix} \mu_{n,0} & \cdots & \mu_{n,n-1}\end{bmatrix} (M_{\mathcal{B}})_n^{-1} \tilde{E}_n^\dagger \right) \left((\check{M}_{\mathcal{B}})_N - \tilde{E}_n (M_{\mathcal{B}})_n^{-1}\tilde{E}_n^\dagger\right)^{-1} Q.$$

As consequence,

$$\mathcal{J}_{\check{P}_{n+N}} Q^{-1} = \left[\langle P_n(x), I_p\rangle_W \quad \cdots \quad \left\langle P_n(x), x^{N-1} I_p\right\rangle_W\right] |(\check{M}_{\mathcal{B}})_{n+N}|^{-1},$$

where $|(\check{M}_{\mathcal{B}})_{n+N}|$ was defined in (3.23). Let $\varepsilon_{n+N} =: \left(\langle P_n(x), I_p\rangle_W \cdots \langle P_n(x), x^{N-1} I_p\rangle_W \right) |(\check{M}_{\mathcal{B}})_{n+N}|^{-1}$. Therefore, we can establish the following connection formula.

Proposition 3.19. *Assuming that $\langle \cdot, \cdot \rangle_W$ is a quasi-definite sesquilinear form, the following connection formula holds*

$$\check{P}_{n+N}(x) = P_n(x)W(x) + \varepsilon_{n+N} \left(\begin{bmatrix} I_p \\ \vdots \\ x^{N-1} I_p \end{bmatrix} - \sum_{k=0}^{n-1} \begin{bmatrix} E_{0,k} \\ \vdots \\ E_{N-1,k} \end{bmatrix} P_k(x)W(x) \right),$$

where

$$E_{l,k} = \left\langle x^l I_p, P_k W(x) \right\rangle_W \|P_k\|_L^{-2}.$$

Proof. Let $\check{P}_{n+N}(x) = \sum_{m\geqslant 0} \sum_{l=0}^{N-1} \check{a}_{l,m}^{[n+N]} x^l W^m(x)$ be the matrix orthogonal polynomial of degree $N+n$ with respect to $\langle \cdot, \cdot \rangle_W$ given in terms of the basis $\mathcal{B}_w$. Since $(P_n(x)W(x))_{n\in\mathbb{N}}$ is a basis of the left module $\mathbb{C}^{p\times p}[x]W(x)$, then there exist matrices $(\gamma_{n,k})_{k=0}^{n-1}$ such that

$$\check{P}_{n+N}(x) - \sum_{l=0}^{N-1} \check{a}_{l,0}^{[n+N]} x^l = P_n(x)W(x) + \sum_{k=0}^{n-1} \gamma_{n,k} P_k(x)W(x).$$

Since for $k = 1, \ldots, n-1$, we have

$$\left\langle \check{P}_{n+N}(x) - \sum_{l=0}^{N-1} \check{a}_{l,0}^{[n+N]} x^l, P_k(x)W(x) \right\rangle_W = - \sum_{l=0}^{N-1} \check{a}_{l,0}^{[n+N]} \left\langle x^l I_p, P_k(x)W(x) \right\rangle_W, \tag{3.28}$$

as well as,

$$\gamma_{n,k} = - \sum_{l=0}^{N-1} \check{a}_{l,0}^{[n+N]} \left\langle x^l I_p, P_k(x)W(x) \right\rangle_W \|P_k\|_L^{-2} = - \sum_{l=0}^{N-1} \check{a}_{l,0}^{[n+N]} E_{l,k}. \tag{3.29}$$

From (3.28) and (3.29)

$$\check{P}_{n+N}(x) = P_n(x)W(x) + \sum_{l=0}^{N-1} \check{a}_{l,0}^{[n+N]} \left(x^l I_p - \sum_{k=0}^{n-1} E_{l,k} P_k(x)W(x) \right), \tag{3.30}$$

and since $\mathcal{J}_{\check{P}_{n+N}} Q^{-1} = \begin{bmatrix} \check{a}_{0,0}^{[n+N]} & \check{a}_{1,0}^{[n+N]} & \cdots & \check{a}_{N-1,0}^{[n+N]} \end{bmatrix}$ we get the result. ∎

Proposition 3.20. *The sequences $(P_n(x))_{n\in\mathbb{N}}$ and $(\check{P}_n(x))_{n\in\mathbb{N}}$ satisfy the following inverse connection formula*

$$PW(x) = \mathcal{M}\check{P}, \tag{3.31}$$

where $\check{P} = [\check{P}_0^\dagger(x), \check{P}_1^\dagger(x), \dots]^\dagger$ and M is a block lower Hessenberg matrix with block entries

$$\beta_{n,k} = \begin{cases} I_p, & \text{if} \quad k = n+N, \\ \left\langle P_n(x)W(x), \sum_{l=0}^{N-1} \check{a}_{l,0}^{[k]} x^l \right\rangle_W \|\check{P}_k\|_W^{-2}, & \text{if} \quad 0 \le k \le (N-1)+n, \\ 0_p, & \text{otherwise}. \end{cases}$$

Proof. Since $(\check{P}_n(x))_{n\in\mathbb{N}}$ is a basis of the left module $\mathbb{C}^{p\times p}[x]$, then there exist matrices $(\beta_{n,k})_{k=0}^{n+N-1}$ such that

$$P_n(x)W(x) = \check{P}_{n+N}(x) + \sum_{k=0}^{n+N-1} \beta_{n,k}\check{P}_k(x).$$

If $k < N$, then

$$\beta_{n,k} = \left\langle P_n(x)W(x), \sum_{l=0}^{N-1} \check{a}_{l,0}^{[k]} x^l \right\rangle_W \|\check{P}_k\|_W^{-2}.$$

On the other hand, if $k \ge N$, using (3.30) we get

$$\begin{aligned} \langle P_nW(x), \check{P}_k(x)\rangle_W &= \left\langle P_nW(x), P_{k-N}(x)W(x) + \sum_{l=0}^{N-1} \check{a}_{l,0}^{[k]} \left(x^l I_p - \sum_{i=0}^{k-N-1} E_{l,i}P_i(x)W(x) \right) \right\rangle_W \\ &= \sum_{l=0}^{N-1} \left(\left\langle P_nW(x), x^l I_p \right\rangle_W - \sum_{i=0}^{k-N-1} \langle P_nW(x), P_i(x)W(x)\rangle_W E_{l,i}^\dagger \right) \check{a}_{l,0}^{[k]\dagger} \\ &= \left\langle P_nW(x), \sum_{l=0}^{N-1} \check{a}_{l,0}^{[k]} x^l \right\rangle_W . \end{aligned}$$

The above implies that for $k = 0, \dots, N+n-1$,

$$\beta_{n,k} = \left\langle P_n(x)W(x), \sum_{l=0}^{N-1} \check{a}_{l,0}^{[k]} x^l \right\rangle_W \|\check{P}_k\|_W^{-2}.$$

■

Remark 3.21. *The matrices ε_{n+N}, $E_{l,k}$, and $\beta_{n,k}$ appearing in Propositions 3.19 and 3.20, which depend on the sesquilinear form $\langle\cdot,\cdot\rangle_W$, can be obtained using only "non-perturbed" data according to Proposition 3.15.*

Since $\langle\cdot,\cdot\rangle_W$ is a sesquilinear form that not necessarily satisfies $\langle xP, Q\rangle_W = \langle P, xQ\rangle_W$ for every $P, Q \in \mathbb{C}^{p\times p}[x]$, then the semi-infinite block matrix $\check{J}$ associated with the multiplication operator by x with respect to the sequence of polynomials $(\check{P}_n(x))_{n\in\mathbb{N}}$ (i.e., $\check{P}x = \check{J}\check{P}$) is a block Hessenberg matrix.

Proposition 3.22. *Let $W(x) = \sum_{j=0}^{N} c_j x^j$, $c_j \in \mathbb{C}^{p\times p}$. If J and $\check{J}$ are the block Jacobi and block Hessenberg matrices associated with the sequences of monic matrix orthogonal polynomials $(P_n(x))_{n\in\mathbb{N}}$, $(\check{P}_n(x))_{n\in\mathbb{N}}$, respectively, then*

$$W_S(J) = \mathcal{M}L, \quad W_{\check{S}}(\check{J}) = L\mathcal{M},$$

where $W_S(x) = \sum_{j=0}^{N}(S\beta_j)S^{-1}x^j$, $W_{\check{S}}(x) = \sum_{j=0}^{N}(\check{S}\beta_j)\check{S}^{-1}x^j$ (see Remark 1.7) and L is the lower triangular block matrix with I_p as diagonal entries and such that $\check{P} = LP$.

Proof. From the hypothesis and (3.31), we get $PW(x) = (\mathcal{M}L)P$ and $\check{P}W(x) = (L\mathcal{M})\check{P}$. On the other hand, taking into account $P = S\chi(x)$ and the properties of the shift matrix, we have

$$\begin{aligned} PW(x) &= \sum_{j=0}^{N}\left((Sc_j)S^{-1}\right)Px^j \\ &= \sum_{j=0}^{N}\left((Sc_j)S^{-1}\right)JPx^{j-1} = \cdots = \sum_{j=0}^{N}\left((Sc_j)S^{-1}\right)J^jP \\ &= W_S(J)P. \end{aligned}$$

Thus $(W_S(J) - \mathcal{M}L)P = \mathbf{0}$ where $\mathbf{0}$ is the semi-infinite matrix of zeros. Since J has Jacobi block structure, then it is easy to see that both $W_T(J)$ and $\mathcal{M}L$ are block Hessenberg matrices with shape

$$\begin{bmatrix} \overbrace{* \quad * \quad \cdots \quad I_p}^{N+n} & & & \\ * \quad * \quad \cdots \quad * & I_p & & \\ * \quad * \quad \cdots \quad * & * & I_p & \\ \cdots \quad \cdots \quad \cdots \quad \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

Since $(P_n(x))_{n\in\mathbb{N}}$ is a basis of the left module $\mathbb{C}^{p\times p}[x]$ we conclude that $W_S(J) - \mathcal{M}L = \mathbf{0}$. The other equation can be obtained in a similar way. ∎

Chapter 4

Christoffel transformations for matrix bi-orthogonal polynomials on the real line and the non-Abelian 2D Toda lattice hierarchy

In this Chapter, we focus our attention on the study of Christoffel transformations for matrix linear functionals. More precisely, given a matrix of linear functionals u and a matrix polynomial $W(x)$ we will deal with the matrix of linear functionals $\hat{u}$ defined as

$$\hat{u} = W(x)u.$$

We will first focus our attention on the existence of matrix bi-orthogonal polynomials with respect to the sesquilinear form $\langle\cdot,\cdot\rangle_{\hat{u}}$ defined from $\hat{u}$, making some assumptions about the matrix polynomial W. Once this is done, the next step will be to find an explicit representation of such bi-orthogonal polynomials in terms of the matrix bi-orthogonal polynomials with respect to the matrix of linear functionals u.

4.1 Connection formulas for Darboux transformations of Christoffel type

Given a monic matrix polynomial $W(x) \in \mathbb{C}^{p\times p}[x]$ of degree N with different zeros $x_1,\dots,x_q$, we consider a new matrix of linear functionals $\hat{u}$ such that

$$u \mapsto \hat{u}(x) := W(x)u$$

and the corresponding perturbed sesquilinear form

$$\langle P(x), Q(x)\rangle_{\hat{u}} = \langle P(x)W(x), Q(x)\rangle_{u}.$$

Recall that the moment block matrix associated with the matrix linear functional u is given by $M = \langle \chi(x), \chi(x) \rangle_u$ (see Remark 1.7). In the same way the moment block matrix

$$\hat{M} =: \langle \chi(x), \chi(x) \rangle_{\hat{u}} = \langle \chi(x) W(x), \chi(x) \rangle_u \tag{4.1}$$

is introduced. Let us assume that the perturbed moment matrix has a Gaussian factorization

$$\hat{M} = \hat{S}_1^{-1} \hat{H} (\hat{S}_2)^{-\dagger},$$

where $\hat{S}_1, \hat{S}_2$ are lower unitriangular block matrices and $\hat{H}$ is a diagonal block matrix

$$\hat{S}_i = \begin{bmatrix} I_p & 0 & 0 & \dots \\ (\hat{S}_i)_{1,0} & I_p & 0 & \cdots \\ (\hat{S}_i)_{2,0} & (\hat{S}_i)_{2,1} & I_p & \ddots \\ & & & \ddots \end{bmatrix}, \qquad \hat{H} = \begin{bmatrix} \hat{H}_0 & 0 & 0 & \cdots \\ 0 & \hat{H}_1 & 0 & \ddots \\ 0 & 0 & \hat{H}_2 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \qquad i = 1, 2.$$

Then, we have the corresponding perturbed bi-orthogonal matrix polynomials

$$\hat{P}^{[i]}(x) = \hat{S}_i \chi(x), \qquad i = 1, 2,$$

with respect to the perturbed sesquilinear form $\langle \cdot, \cdot \rangle_{\hat{u}}$.

Remark 4.1. *For monic matrix polynomial perturbations and perturbations with a matrix polynomial with nonsingular leading coefficients the analysis of these problems are equivalent. Indeed, if instead of a monic matrix polynomial we have a matrix polynomial $\tilde{W}(x) = A_N x^N + \cdots + A_0$ with a nonsingular leading coefficient, $\det A_N \neq 0$, $\tilde{W}(x) = A_N W(x)$, where W is monic. The moment matrices are related by $\tilde{M} = A_N \hat{M}$ and, moreover, $\tilde{S}_1 = A_N \hat{S}_1 (A_N)^{-1}$, $\tilde{H} = A_N \hat{H}$, $\tilde{S}_2 = \hat{S}_2$, and $\tilde{P}_k^{[1]}(x) = A_N P_k^{[1]}(x)(A_N)^{-1}$ as well as $\tilde{P}_k^{[2]}(x) = \hat{P}_k^{[2]}(x)$.*

4.1.1 Connection formulas for bi-orthogonal polynomials

Proposition 4.2. *The moment matrix M and the perturbed moment matrix $\hat{M}$ satisfy*

$$\hat{M} = W(\Lambda) M.$$

Proof. This is a direct consequence of (4.1). ■

Definition 4.3. *Let us introduce the following semi-infinite matrices*

$$\omega^{[1]} := \hat{S}_1 W(\Lambda) S_1^{-1}, \qquad \omega^{[2]} := (S_2 \hat{S}_2^{-1})^{\dagger},$$

which we call resolvent or connection matrices.

Proposition 4.4 (Connection formulas)**.** *Perturbed and non perturbed bi-orthogonal polynomials satisfy the following linear connection formulas*

$$\omega^{[1]}P^{[1]}(x) = \hat{P}^{[1]}(x)W(x), \tag{4.2}$$

$$P^{[2]}(x) = (\omega^{[2]})^{\dagger}\hat{P}^{[2]}(x). \tag{4.3}$$

Proposition 4.5. *The following relations hold*

$$\hat{H}\omega^{[2]} = \omega^{[1]}H.$$

Proof. From Proposition 1.3, Proposition 4.2, and the Gauss-Borel factorization we get

$$\hat{S}_1^{-1}\hat{H}\hat{S}_2^{-\dagger} = W(\Lambda)S_1^{-1}HS_2^{-\dagger},$$

so that

$$\hat{H}(S_2\hat{S}_2^{-1})^{\dagger} = \hat{S}_1 W(\Lambda)S_1^{-1}H$$

and the result follows. ∎

From the above two results we easily get

Proposition 4.6. *The resolvent matrix ω is a band upper triangular block matrix with all block superdiagonals above the N-th one equal to zero.*

$$\omega^{[1]} = \begin{bmatrix} \omega^{[1]}_{0,0} & \omega^{[1]}_{0,1} & \omega^{[1]}_{0,2} & \dots & \omega^{[1]}_{0,N-1} & I_p & 0 & 0 & \dots \\ 0 & \omega^{[1]}_{1,1} & \omega^{[1]}_{1,2} & \dots & \omega^{[1]}_{1,N-1} & \omega^{[1]}_{1,N} & I_p & 0 & \dots \\ 0 & 0 & \omega^{[1]}_{2,2} & \dots & \omega^{[1]}_{2,N-1} & \omega^{[1]}_{2,N} & \omega^{[1]}_{2,N+1} & I_p & \ddots \\ & \ddots & \ddots & \ddots & & & & \ddots & \ddots \end{bmatrix}$$

with

$$\omega^{[1]}_{k,k} = \hat{H}_k(H_k)^{-1}. \tag{4.4}$$

Remark 4.7. *Notice that* (4.2) *and* (4.3) *implies that*

$$\hat{P}^{[1]}_n(x)W(x) = \sum_{k=n}^{n+N} \omega^{[1]}_{n,k}P_k(x),$$

$$P^{[2]}_n(x) = \sum_{k=\max\{n-N,0\}}^{n} \left(H_n^{\dagger}\omega^{[1]\dagger}_{k,n}\hat{H}_k^{-\dagger}\right)\hat{P}^{[2]}_k(x).$$

4.1.2 Connection formulas for the Christoffel–Darboux kernel

In order to relate the perturbed and non perturbed kernel matrix polynomials (see Definition 1.66) let us introduce the following truncation of the connection matrix ω.

Definition 4.8. *We introduce the lower unitriangular matrix* $\omega_{(n,N)} \in \mathbb{C}^{Np\times Np}$

$$\omega_{(n,N)} := \begin{cases} \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & \vdots & & \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \omega^{[1]}_{0,n+1} & \dots & \omega_{0,N-1} & I_p & \ddots & 0 \\ \vdots & & & & \ddots & \vdots \\ \omega^{[1]}_{n,n+1} & \dots & & & \omega^{[1]}_{n,n+N-1} & I_p \end{bmatrix}, & n<N, \\ \begin{bmatrix} I_p & 0 & \dots & 0 & 0 \\ \omega^{[1]}_{n-N+2,n+1} & I_p & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & I_p & 0 \\ \omega^{[1]}_{n,n+1} & \dots & & \omega^{[1]}_{n,n+N-1} & I_p \end{bmatrix}, & n\geq N, \end{cases}$$

and the diagonal block matrix

$$\hat{H}_{n,N} = \operatorname{diag}(\hat{H}_{n-N+1},\dots,\hat{H}_n).$$

Then, we can state

Theorem 4.9. *The perturbed and original Christoffel-Darboux kernels are related by the following connection formulas*

$$K_n(x,y) + \left[(\hat{P}^{[2]}_{n-N+1}(y))^\dagger,\dots,(\hat{P}^{[2]}_n(y))^\dagger\right](\hat{H}_{(n,N)})^{-1}\omega_{(n,N)}\begin{bmatrix} P^{[1]}_{n+1}(x) \\ \vdots \\ P^{[1]}_{n+N}(x)\end{bmatrix} = \hat{K}_n(x,y)W(x),$$

where, by convention, $\hat{P}^{[2]}_j(x) = 0$ *whenever* $j<0$.

Proof. Consider the truncation

$$(\omega^{[2]})_{[n+1]} := \begin{bmatrix} I_p & \dots & \omega^{[2]}_{0,N-1} & \omega^{[2]}_{0,N} & 0 & 0 & \dots & 0 \\ 0 & I_p & \dots & \omega^{[2]}_{1,N+1} & \omega^{[2]}_{1,N+2} & 0 & \dots & 0 \\ \vdots & & \ddots & & & \ddots & \ddots & \\ 0 & 0 & \cdots & I_p & \cdots & & & \omega^{[2]}_{n-N,n} \\ \vdots & \vdots & & & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & & & \ddots & & \vdots \\ 0 & 0 & & & & & I_p & \omega^{[2]}_{n-1,n} \\ 0 & 0 & & & & & 0 & I_p \end{bmatrix}.$$

Rewriting (4.3) as $(P^{[2]}(y))^{\dagger} = (\hat{P}^{[2]}(y))^{\dagger}\omega^{[2]}$, then $\big((\hat{P}^{[2]}(y))_{[n+1]}\big)^{\dagger}(\omega^{[2]})_{[n+1]} = \big((P^{[2]}(y))_{[n+1]}\big)^{\dagger}$ holds for the $n+1$-th truncations of $P^{[2]}(y))$ and $\hat{P}^{[2]}(y)$. Therefore,

$$\begin{aligned}\big((\hat{P}^{[2]}(y))_{[n+1]}\big)^{\dagger}(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]} =& \big((P^{[2]}(y))_{[n+1]}\big)^{\dagger}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]} \\ =& K_n(x,y).\end{aligned}$$

Now, we consider $(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]}$ and from Proposition 4.5

$$(\omega^{[2]})_{[n+1]}\big(H_{[n+1]}\big)^{-1} = (\hat{H}_{[n+1]})^{-1}(\omega^{[1]})_{[n+1]}$$

we deduce

$$(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]} = (\hat{H}_{[n+1]})^{-1}(\omega^{[1]})_{[n+1]}(P^{[1]}(x))_{[n+1]}.$$

Observe also that

$$(\omega^{[1]})_{[n+1]}(P^{[1]}(x))_{[n+1]} = (\omega^{[1]}P^{[1]}(x))_{[n+1]} - \begin{bmatrix} 0_{(n-N)p\times p} \\ V_N(x) \end{bmatrix}$$

with

$$V_N(x) = \omega_{(n,N)} \begin{bmatrix} P^{[1]}_{n+1}(x) \\ \vdots \\ P^{[1]}_{n+N}(x) \end{bmatrix}.$$

Hence, recalling (4.2) we get

$$(\omega^{[1]})_{[n+1]}(P^{[1]}(x))_{[n+1]} = (\hat{P}^{[1]}(x))_{[n+1]}W(x) - \begin{bmatrix} 0_{(n-N)p\times p} \\ V_N(x) \end{bmatrix},$$

and, consequently,

$$
\begin{aligned}
&\big((\hat{P}^{[2]}(y))_{[n+1]}\big)^{\dagger}(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]}\\
&\quad=\big((\hat{P}^{[2]}(y))_{[n+1]}\big)^{\dagger}(\hat{H}_{[n+1]})^{-1}(\hat{P}^{[1]}(x))_{[n+1]}W(x)-\big((\hat{P}^{[2]}(y))_{[n+1]}\big)^{\dagger}(\hat{H}_{[n+1]})^{-1}\begin{bmatrix}0_{(n-N)p\times p}\\ V_N(x)\end{bmatrix}\\
&\quad=\hat{K}_n(x,y)W(x)-\big((\hat{P}^{[2]}(y))_{[n+1]}\big)^{\dagger}(\hat{H}_{[n+1]})^{-1}\begin{bmatrix}0_{(n-N)p\times p}\\ V_N(x)\end{bmatrix}.
\end{aligned}
$$

■

4.2 Monic matrix polynomial perturbations

In this section we study perturbations by a monic matrix polynomial $W(x)$, which is equivalent to the study of perturbations by a matrix polynomials with nonsingular leading coefficients as we have pointed out above. Using the theory presented in Chapter 1, we are able to extend the celebrated Christoffel formula to this context.

4.2.1 The Christoffel formula for matrix bi-orthogonal polynomials

We are now ready to show how the perturbed set of matrix bi-orthogonal polynomials $(\hat{P}_n^{[1]}(x),\hat{P}_n^{[2]}(x))_{n\in\mathbb{N}}$ is related to the original set $(P_n^{[1]}(x),P_n^{[2]}(x))_{n\in\mathbb{N}}$.

Proposition 4.10. *For each $i\in\{1,\dots,q\}$, let $r_j^{(i)}(x)$, $j=1,\dots,s_i$, be the adapted root polynomials of the monic matrix polynomial $W(x)$ corresponding to the eigenvalue x_i (see Definition 1.26). Then*

$$
\omega_{k,k}^{[1]}\frac{\mathrm{d}^l(P_k^{[1]}r_j^{(i)})}{\mathrm{d}x^l}\bigg|_{x=x_i}+\cdots+\omega_{k,k+N-1}^{[1]}\frac{\mathrm{d}^l(P_{k+N-1}^{[1]}r_j^{(i)})}{\mathrm{d}x^l}\bigg|_{x=x_i}=-\frac{\mathrm{d}^l(P_{k+N}^{[1]}r_j^{(i)})}{\mathrm{d}x^l}\bigg|_{x=x_i},\tag{4.5}
$$

for $l=0,\dots,\kappa_j^{(i)}-1$. Moreover (see Definition 1.35)

$$
-\mathcal{J}_{P_{k+N}}(x_i)=\omega_{k,k+N-1}^{[1]}\mathcal{J}_{P_{k+N-1}}(x_i)+\cdots+\omega_{k,k}^{[1]}\mathcal{J}_{P_k}(x_i).\tag{4.6}
$$

Proof. From (4.2) we get

$$
\omega_{k,k}^{[1]}P_k^{[1]}(x)+\cdots+\omega_{k,k+N-1}^{[1]}P_{k+N-1}^{[1]}(x)+P_{k+N}(x)=\hat{P}_k^{[1]}(x)W(x).
$$

Now, according to Proposition 1.33 we have

$$
\frac{\mathrm{d}^l}{\mathrm{d}x^l}\bigg|_{x=x_i}\big(\hat{P}_k^{[1]}W(x)r_j^{(i)}(x)\big)=\sum_{s=0}^{l}\binom{l}{s}\frac{\mathrm{d}^{l-s}\hat{P}_k^{[1]}(x)}{\mathrm{d}x^{l-s}}\bigg|_{x=x_i}\frac{\mathrm{d}^s(W(x)r_j^{(i)}(x))}{\mathrm{d}x^s}\bigg|_{x=x_i}=0\tag{4.7}
$$

for $l=0,\dots,\kappa_j^{(i)}-1$ and $j=1,\dots,s_i$. (4.6) is an immediate consequence of (4.5). ■

Recall that $\sum_{j=1}^{s_i}\kappa_j^{(i)}=\alpha_i$ and $Np=\sum_{i=1}^{q}\alpha_i$, $q=\#\sigma(W)$.

Theorem 4.11 (The Christoffel formula for matrix bi-orthogonal polynomials)**.** *The perturbed set of matrix bi-orthogonal polynomials $(\hat{P}_n^{[1]}(x),\hat{P}_n^{[2]}(x))_{n\in\mathbb{N}}$ can be written as the following last quasi-determinant*

$$\hat{P}_n^{[1]}(x)W(x)=\Theta_*\left[\begin{array}{ccc|c}\mathcal{J}_{P_n^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_n^{[1]}}(x_q) & P_n^{[1]}(x)\\ \vdots & & \vdots & \vdots\\ \mathcal{J}_{P_{n+N-1}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N-1}^{[1]}}(x_q) & P_{n+N-1}^{[1]}(x)\\ \hline \mathcal{J}_{P_{n+N}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N}^{[1]}}(x_q) & P_{n+N}^{[1]}(x)\end{array}\right]=\Theta_*\left[\begin{array}{c|c}\mathcal{J}_{P_n^{[1]}} & P_n^{[1]}(x)\\ \vdots & \vdots\\ \mathcal{J}_{P_{n+N-1}^{[1]}} & P_{n+N-1}^{[1]}(x)\\ \hline \mathcal{J}_{P_{n+N}^{[1]}} & P_{n+N}^{[1]}(x)\end{array}\right],\tag{4.8}$$

$$(\hat{P}_n^{[2]}(y))^{\dagger}(\hat{H}_n)^{-1}=\Theta_*\left[\begin{array}{ccc|c}\mathcal{J}_{P_{n+1}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+1}^{[1]}}(x_q) & 0\\ \vdots & & \vdots & \vdots\\ & & & 0\\ \mathcal{J}_{P_{n+N}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N}^{[1]}}(x_q) & I_p\\ \hline \mathcal{J}_{K_n}(x_1,y) & \cdots & \mathcal{J}_{K_n}(x_q,y) & 0\end{array}\right]=\Theta_*\left[\begin{array}{c|c}\mathcal{J}_{P_{n+1}^{[1]}} & 0\\ \vdots & \vdots\\ & 0\\ \mathcal{J}_{P_{n+N}^{[1]}} & I_p\\ \hline \mathcal{J}_{K_n}(y) & 0\end{array}\right].\tag{4.9}$$

Moreover, the new matrix squared norms are

$$\hat{H}_n=\Theta_*\left[\begin{array}{ccc|c}\mathcal{J}_{P_n^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_n^{[1]}}(x_q) & H_k\\ & & & 0\\ \vdots & & \vdots & \vdots\\ \mathcal{J}_{P_{n+N-1}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N-1}^{[1]}}(x_q) & 0\\ \hline \mathcal{J}_{P_{n+N}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N}^{[1]}}(x_q) & 0\end{array}\right]=\Theta_*\left[\begin{array}{c|c}\mathcal{J}_{P_n^{[1]}} & H_k\\ & 0\\ \vdots & \vdots\\ \mathcal{J}_{P_{n+N-1}^{[1]}} & 0\\ \hline \mathcal{J}_{P_{n+N}^{[1]}} & 0\end{array}\right].\tag{4.10}$$

Proof. We assume that $P_j^{[2]}(x)=0$ whenever $j<0$. To prove (4.8), notice that from (4.5) for the rows of the connection matrix we get

$$[\omega_{n,n}^{[1]},\dots,\omega_{n,n+N-1}^{[1]}]=-\begin{bmatrix}\mathcal{J}_{P_{n+N}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N}^{[1]}}(x_q)\end{bmatrix}\begin{bmatrix}\mathcal{J}_{P_n^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_n^{[1]}}(x_q)\\ \vdots & & \vdots\\ \mathcal{J}_{P_{n+N-1}^{[1]}}(x_1) & \cdots & \mathcal{J}_{P_{n+N-1}^{[1]}}(x_q)\end{bmatrix}^{-1}.\tag{4.11}$$

Now, using (4.2)

$$[\omega_{n,n}^{[1]},\dots,\omega_{n,n+N-1}^{[1]}]\begin{bmatrix}P_n^{[1]}(x)\\ \vdots\\ P_{n+N-1}^{[1]}(x)\end{bmatrix}+P_{n+N}^{[1]}(x)=\hat{P}_n^{[1]}(x)W(x),$$

and (4.8) follows immediately.
To deduce (4.9) for $n \geq N$ notice that Theorem 4.9 together with (4.7) yields

$$\mathcal{J}_{K_n}(x_i,y)+\begin{bmatrix}(\hat{P}^{[2]}_{n-N+1}(y))^{\dagger} & \dots & (\hat{P}^{[2]}_n(y))^{\dagger}\end{bmatrix}(\hat{H}_{(n,N)})^{-1}\omega_{(n,N)}\begin{bmatrix}\mathcal{J}_{P^{[1]}_{n+1}}(x_i)\\ \vdots \\ \mathcal{J}_{P^{[1]}_{n+N}}(x_i)\end{bmatrix}=0$$

for $i=1,\cdots,q$. If we arrange these equations in a matrix form

$$\begin{bmatrix}\mathcal{J}_{K_n}(x_1,y) & \cdots & \mathcal{J}_{K_n}(x_q,y)\end{bmatrix}+\begin{bmatrix}(\hat{P}^{[2]}_{n-N+1}(y))^{\dagger} & \dots & (\hat{P}^{[2]}_n(y))^{\dagger}\end{bmatrix}(\hat{H}_{(n,N)})^{-1}\omega_{(n,N)}\begin{bmatrix}\mathcal{J}_{P^{[1]}_{n+1}}(x_1) & \cdots & \mathcal{J}_{P^{[1]}_{n+1}}(x_q)\\ \vdots & & \vdots\\ \mathcal{J}_{P^{[1]}_{n+N}}(x_1) & \cdots & \mathcal{J}_{P^{[1]}_{n+N}}(x_q)\end{bmatrix}=0.$$

Therefore, assuming that the above matrix is nonsingular , we get

$$\begin{bmatrix}(\hat{P}^{[2]}_{n-N+1}(y))^{\dagger} & \dots & (\hat{P}^{[2]}_n(y))^{\dagger}\end{bmatrix}(\hat{H}_{(n,N)})^{-1}\omega_{(n,N)}=-\begin{bmatrix}\mathcal{J}_{K_n}(x_1,y) & \cdots & \mathcal{J}_{K_n}(x_q,y)\end{bmatrix}\begin{bmatrix}\mathcal{J}_{P^{[1]}_{n+1}}(x_1) & \cdots & \mathcal{J}_{P^{[1]}_{n+1}}(x_q)\\ \vdots & & \vdots\\ \mathcal{J}_{P^{[1]}_{n+N}}(x_1) & \cdots & \mathcal{J}_{P^{[1]}_{n+N}}(x_q)\end{bmatrix}^{-1},$$

which, in particular, gives

$$(\hat{P}^{[2]}_n(y))^{\dagger}(\hat{H}_n)^{-1}=-\begin{bmatrix}\mathcal{J}_{K_n}(x_1,y) & \cdots & \mathcal{J}_{K_n}(x_q,y)\end{bmatrix}\begin{bmatrix}\mathcal{J}_{P^{[1]}_{n+1}}(x_1) & \cdots & \mathcal{J}_{P^{[1]}_{n+1}}(x_q)\\ \vdots & & \vdots\\ \mathcal{J}_{P^{[1]}_{n+N}}(x_1) & \cdots & \mathcal{J}_{P^{[1]}_{n+N}}(x_q)\end{bmatrix}^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ I_p\end{bmatrix}.$$

Finally, (4.10) is a consequence of (4.4) and (4.11). ■

4.2.2 Degree one monic matrix polynomial perturbations

Let us illustrate the situation with the most simple case of a perturbation with a monic polynomial matrix of degree one.

Proposition 4.12 (Degree one Christoffel formula)**.** *If $W(x)=I_px-A$ and $\det P^{[1]}_n(A)\neq 0$ for $n\in\mathbb{N}$, then the Christoffel formulas can be written as*

$$\begin{aligned}\hat{P}^{[1]}_n(x)(I_px-A)&=\Theta_*\begin{bmatrix}P^{[1]}_n(A) & P^{[1]}_n(x)\\ P^{[1]}_{n+1}(A) & P^{[1]}_{n+1}(x)\end{bmatrix}\\ &=P^{[1]}_{n+1}(x)-P^{[1]}_{n+1}(A)\left[P^{[1]}_n(A)\right]^{-1}P^{[1]}_n(x),\end{aligned}\qquad \begin{aligned}(\hat{P}^{[2]}_n(y))^{\dagger}(\hat{H}_n)^{-1}&=\Theta_*\begin{bmatrix}P^{[1]}_{n+1}(A) & I\\ K_n(A,y) & 0\end{bmatrix}\\ &=-K_n(A,y)\left[P^{[1]}_n(A)\right]^{-1}.\end{aligned}$$

For the perturbed matrix squared norms we have

$$\hat{H}_n=\Theta_*\begin{bmatrix}P^{[1]}(A) & H_n\\ P^{[n+1]}(A) & 0\end{bmatrix}=-P^{[1]}_{n+1}(A)\left[P^{[1]}_n(A)\right]^{-1}H_n.$$

Proof. According to (4.2) and Theorem 4.9

$$\omega_{n,n}P_n^{[1]}(x)+P_{n+1}^{[1]}(x)=\hat{P}_n^{[1]}(x)(I_px-A),\quad K_n(x,y)+(\hat{P}_n^{[2]}(y))^{\dagger}\hat{H}_n^{-1}P_{n+1}^{[1]}(x)=\hat{K}_n(x,y)(I_px-A). \tag{4.12}$$

On the other hand, using Lemma 1.37, we have that $XJ-AX=0$. Thus for any matrix polynomial $R(x)=\sum_{k=0}^{n}R_kx^k$ its spectral jet vector with respect to $W(x)$ is the $p\times p$ matrix

$$\mathcal{J}_R=\sum_{k=0}^{n}R_k\mathcal{J}_{x^k}=\sum_{k=0}^{n}R_kA^kX=R(A)X.$$

In particular, (4.12) reads

$$\omega_{n,n}P_n^{[1]}(A)X+P_{n+1}^{[1]}(A)X=0_p,\qquad K_n(A,y)X+(\hat{P}_n^{[2]}(y))^{\dagger}(\hat{H}_n)^{-1}P_{n+1}^{[1]}(A)X=0_p,$$

and since that in this case X is a nonsingular matrix, the result follows. ■

Next, we will illustrate the Christoffel formula in the matrix orthogonal polynomial framework in a simple situation. We will study what is the effect of the Christoffel transformation on a positive Borel scalar measure $\mathrm{d}\mu(x)$, thus the perturbed matrix of measures is $(I_2x-A)\,\mathrm{d}\mu(x)$. The perturbed monic orthogonal polynomials will be expressed, see Proposition 4.12,

$$\begin{aligned}\hat{P}_n^{[1]}(x)&=\Big(I_2p_{n+1}(x)-p_{n+1}(A)\big(p_n(A)\big)^{-1}p_n(x)\Big)(I_2x-A)^{-1},\\ \hat{P}_n^{[2]}(x)(\hat{H}_n)^{-1}&=-K_n(x,A)\big[P_n^{[1]}(A)\big]^{-1},\end{aligned}$$

where $p_n(x)$, $K_n(x,y)$, are the scalar orthogonal polynomials and kernel polynomials associated with the original real scalar positive Borel measure $\mathrm{d}\mu(x)$, respectively. Observe that despite starting with a set of orthogonal polynomials the perturbation generates a set of bi-orthogonal matrix polynomials. As the original measure is scalar, if we ensure that $A=A^{\dagger}$ is Hermitian, then we will get $\hat{P}_n(x):=P_n^{[1]}(x)=P_n^{[2]}(x)$, a new set of orthogonal matrix polynomials. However, this will be a very trivial situation. Indeed,

Proposition 4.13. *The matrix orthogonal polynomials $(\hat{P}_n(x))_{n\in\mathbb{N}}$ with respect to the matrix of measures $(I_px-A)\,\mathrm{d}\mu(x)$, where $A=A^{\dagger}$ is Hermitian and $\mathrm{d}\mu$ is a positive Borel scalar measure, are similar to diagonal matrix orthogonal polynomials.*

Proof. Being A a Hermitian matrix it will always be unitarily diagonalizable, i.e.

$$A=QDQ^{\dagger},$$

where Q is an unitary matrix $Q^{\dagger}=Q^{-1}$ and $D=\operatorname{diag}(x_1,\dots,x_p)$, is a diagonal matrix that collects the eigenvalues, not necessarily different, of A. At the end, the new orthogonal polynomials will be

$$\hat{P}_n(x)=Q\begin{bmatrix}\frac{p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1} & 0 & \dots & 0\\ 0 & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_2)}{p_n(x_2)}p_n(x)}{x-x_2} & & 0\\ \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_p)}{p_n(x_p)}p_n(x)}{x-x_p}\end{bmatrix}Q^{\dagger}, \tag{4.13}$$

and the result follows. ■

Thus, we have a diagonal bunch of elementary Darboux transformations of the original scalar orthogonal polynomials associated with the scalar measure $\mathrm{d}\mu$. This situation reappears even when the matrix is not symmetric but diagonalizable, since the perturbed matrix orthogonal polynomials will be similar to a similar bunch of elementary Darboux transformations of the original scalar orthogonal polynomials.

$$\hat{P}^{[1]}_n(x)=Q\begin{bmatrix}\frac{p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1} & 0 & \dots & 0\\ 0 & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_2)}{p_n(x_2)}p_n(x)}{x-x_2} & & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_p)}{p_n(x_p)}p_n(x)}{x-x_p}\end{bmatrix}Q^{-1},$$

$$\hat{P}^{[2]}_n(x)(\hat{H}_n)^{-1}=-Q\begin{bmatrix}\frac{K_n(x,x_1)}{p_n(x_1)} & 0 & \dots & 0\\ 0 & \frac{K_n(x,x_2)}{p_n(x_2)} & & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & & \frac{K_n(x,x_p)}{p_n(x_p)}\end{bmatrix}Q^{-1},$$

where Q does not need to be an orthogonal matrix.

If the matrix is not diagonalizable and has nontrivial Jordan blocks the situation is different. Let us explore this case when $p=2$. Indeed, we consider

$$W(x)=I_2x-A,$$

with

$$A=M\begin{bmatrix}x_1 & 1\\ 0 & x_1\end{bmatrix}M^{-1}.$$

Now we have only one eigenvalue $\sigma(A)=\{x_1\}$, with a Jordan chain of length 2. Thus, there is a basis $\{v_1,v_2\}$ of $\mathbb{C}^2$, $(A-x_1)v_1=v_2$, $(A-x_1)v_2=0$, with $v_i=[v_{i,1}\quad v_{i,2}]^\top$, $i\in\{1,2\}$, such that

$$M=\begin{bmatrix} v_{1,1} & v_{2,1}\\ v_{1,2} & v_{2,2}\end{bmatrix}.$$

Therefore,

$$p_{n+1}(A)\big(p_n(A)\big)^{-1}=M\begin{bmatrix}\frac{p_{n+1}(x_1)}{p_n(x_1)} & \frac{W(p_n,p_{n+1})(x_1)}{(p_n(x_1))^2}\\ 0 & \frac{p_{n+1}(x_1)}{p_n(x_1)}\end{bmatrix}M^{-1},\quad (I_2x-A)^{-1}=M\begin{bmatrix}\frac{1}{x-x_1} & \frac{1}{(x-x_1)^2}\\ 0 & \frac{1}{x-x_1}\end{bmatrix}M^{-1},$$

where

$$W(p_n,p_{n+1})(x)=p_n(x)p'_{n+1}(x)-p_{n+1}(x)p'_n(x)$$

is the Wronskian of two consecutive orthogonal polynomials. Hence,

$$\begin{aligned}\hat{P}^{[1]}_n(x)&=M\begin{bmatrix}\frac{p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1} & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)-\frac{W(p_n,p_{n+1})(x_1)}{(p_n(x_1))^2}(x-x_1)p_n(x)}{(x-x_1)^2}\\ 0 & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1}\end{bmatrix}M^{-1}, \qquad (4.14)\\ \hat{P}^{[2]}_n(x)(\hat{H}_n)^{-1}&=-M\begin{bmatrix}\frac{K_n(x,x_1)}{p_n(x_1)} & -\frac{K_n(x,x_1)}{p_n(x_1)}p'_n(x_1)+\frac{1}{p_n(x_1)}\frac{\partial K_n(x,y)}{\partial y}\Big|_{y=x_1}\\ 0 & \frac{K_n(x,x_1)}{p_n(x_1)}\end{bmatrix}M^{-1}.\end{aligned}$$

Observe that the polynomials

$$p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x),\qquad p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)-\frac{W(p_n,p_{n+1})(x_1)}{(p_n(x_1))^2}(x-x_1)p_n(x)$$

have a zero at $x=x_1$ with multiplicities 1 and 2, respectively.

4.2.3 Examples

4.14 Example . In [81] the authors define the notion of a classical pair $\{w(x),D\}$, where $w(x)$ is a symmetric matrix valued weight function and D is a second order linear ordinary differential operator. In that paper a weight function is said to be classical if there exists a second order linear ordinary differential operator D with matrix valued polynomial coefficients $A_j(t)$, $\deg A_j\leq j$, of the form $D=A_2(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}+A_1(x)\frac{\mathrm{d}}{\mathrm{d}x}+A_0(x)$, such that $\langle DP,Q\rangle=\langle P,DQ\rangle$ for every matrix valued polynomial functions $P(x)$ and $Q(x)$. Then, the pair $\{w,D\}$ is said to be *a classical pair*. In example 5.1 in [81] the authors deal with a family of Jacobi type classical pairs that contains, up to equivalence, all classical pairs of size two where $w(x)=x^\alpha(1-x)^\beta F(x)$, with $\alpha,\beta>-1$ and $0<x<1$, and such that $F(x)$ is a matrix polynomial of degree one, and which are irreducible (in the sense that they are not equivalent to a direct sum of classical pairs of size one). As we will show, they are a direct sum of orthogonal polynomials of size 1 generated by two degree one Christoffel transformations of the scalar Jacobi polynomials with zeros at $x=0$ and $x=1$. Thus, we are faced with two scalar monic Jacobi polynomials with each of the two parameters α and β shifted by one, respectively. In [141] an analysis of the reducibility of matrix weights is given. In particular, in Example 2.4 the case $\alpha=\beta$ is analyzed. We must stress that, as was pointed in [81], reducibility of the matrix of weights $w(x)$ does not imply the reducibility of the classical pair $\{w(x),D\}$. Indeed, despite that the matrix of weights in this example is reducible, the corresponding second order linear differential operator is not.

The classical pair $\{w(x)=x^{\alpha}(1-x)^{\beta}F(x),D\}$ is given by

$$F(x)=F_1x+F_0, \qquad F_1=\begin{bmatrix} 0 & -a \\ -a & \frac{\beta-\alpha}{\alpha+1}a \end{bmatrix}, \qquad F_0=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \text{with } a=\frac{\alpha+\beta+2}{\alpha+1},$$

and a second order matrix linear ordinary differential operator

$$D=x(1-x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}+(X-xU)\frac{\mathrm{d}}{\mathrm{d}x}+V$$

where U,V,X are constant matrices depending on a parameter u. The sequence of orthogonal polynomials $\big(\tilde{\hat{P}}_n^{(\alpha,\beta)}(x)\big)_{n\in\mathbb{N}}$ associated with the classical pair is not given in [81]. Here an explicit representation of $\tilde{\hat{P}}_n^{(\alpha,\beta)}(x)$ using Darboux transformations is deduced. Indeed, we consider that we have an initial alternative Jacobi measure $\mathrm{d}\mu(x)=x^{\alpha}(1-x)^{\beta}I_2$, with $\alpha,\beta>-1$ and $0<x<1$, which is perturbed by a matrix polynomial F of degree one . This matrix polynomial is not monic but its leading coefficient is nonsingular and we can write

$$F(x)=F_1W(x), \qquad W(x)=I_2x-A, \qquad A:=-F_1^{-1}F_0=\frac{1}{a}\begin{bmatrix} \frac{\beta+1}{\alpha+1} & \frac{\beta+1}{\alpha+1} \\ 1 & 1 \end{bmatrix},$$

in terms of a degree one monic matrix polynomial $W(x)$. The matrix A has two different eigenvalues $\sigma(A)=\{0,1\}$ with corresponding eigenvectors $[1,-1]^{\top}$ and $\left[\frac{\beta+1}{\alpha+1},1\right]^{\top}$, as well as the matrix $M:=\begin{bmatrix} 1 & \frac{\beta+1}{\alpha+1} \\ -1 & 1 \end{bmatrix}$ allows to write $A=M\operatorname{diag}(0,1)M^{-1}$.
Let us remember, as was noticed in Remark 4.1, that from the monic orthogonal polynomials $\hat{P}_n^{(\alpha,\beta),[1]}(x)$ with respect to W, we get

$$\tilde{\hat{P}}_n^{(\alpha,\beta)}(x)=F_1\hat{P}_n^{(\alpha,\beta),[1]}(x)F_1^{-1},$$

which are the monic orthogonal polynomials with respect to $w(x)$. Since the matrix of measures $F(x)\,\mathrm{d}\mu(x)$ is symmetric, the bi-orthogonality collapses to orthogonality and the super-indexes [1] and [2] can be omitted. We will do the same with $\hat{P}_n^{(\alpha,\beta),[1]}=\hat{P}_n^{(\alpha,\beta)}$.

Following [32, 33] we conclude that the set of monic matrix orthogonal polynomials $(P_n^{(\alpha,\beta)}(x))_{n\in\mathbb{N}}$ with respect to $\mathrm{d}\mu(x)$ is $P_n^{(\alpha,\beta)}(x)=p_n^{(\alpha,\beta)}(x)I_2$,[1] with the alternative Jacobi polynomials $p_n^{(\alpha,\beta)}(x)$ given by

$$p_n^{(\alpha,\beta)}(x)=\frac{1}{S_n(\alpha,\beta)}\sum_{k=0}^{n}\binom{n+\alpha}{n-k}\binom{n+\beta}{k}x^{n-k}(x-1)^k,\quad \text{where}\quad S_n(\alpha,\beta)=\binom{2n+\beta+\alpha}{n}.$$

On the other hand,

$$p_n^{(\alpha,\beta)}(0)=\frac{1}{S_n(\alpha,\beta)}\binom{n+\alpha}{n}, \qquad p_n^{(\alpha,\beta)}(1)=\frac{1}{S_n(\alpha,\beta)}\binom{n+\beta}{n},$$

so that

$$\frac{p_{n+1}^{(\alpha,\beta)}(0)}{p_n^{(\alpha,\beta)}(0)}=(n+1+\alpha)\rho_n^{(\alpha,\beta)}, \qquad \frac{p_{n+1}^{(\alpha,\beta)}(1)}{p_n^{(\alpha,\beta)}(1)}=(n+1+\beta)\rho_n^{(\alpha,\beta)},$$

where

$$\rho_n^{(\alpha,\beta)}:=\frac{(n+\beta+\alpha+1)}{(2n+\beta+\alpha+2)(2n+\beta+\alpha+1)}.$$

From (4.13) we conclude

$$\hat{P}_n^{(\alpha,\beta)}(x)=M\begin{bmatrix}\frac{p_{n+1}^{(\alpha,\beta)}(x)-(n+1+\alpha)\rho_n^{(\alpha,\beta)}p_n^{(\alpha,\beta)}(x)}{x} & 0\\ 0 & \frac{p_{n+1}^{(\alpha,\beta)}(x)-(n+1+\beta)\rho_n^{(\alpha,\beta)}p_n^{(\alpha,\beta)}(x)}{x-1}\end{bmatrix}M^{-1}.$$

However, let us notice that these two Darboux transformations correspond to the following transformations of the Jacobi measure

$$x^{\alpha}(x-1)^{\beta}\mapsto x(x^{\alpha}(x-1)^{\beta})=x^{\alpha+1}(x-1)^{\beta},\quad x^{\alpha}(x-1)^{\beta}\mapsto(x-1)(x^{\alpha}(x-1)^{\beta})=x^{\alpha}(x-1)^{\beta+1},$$

i.e. the transformations correspond to the shifts $\alpha\mapsto\alpha+1$ and $\beta\mapsto\beta+1$, respectively. Consequently,

$$\hat{P}_n^{(\alpha,\beta)}(x)=M\begin{bmatrix}p_n^{(\alpha+1,\beta)}(x) & 0\\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}M^{-1}.$$

For the matrix

$$\tilde{M}:=\begin{bmatrix}-1 & 1\\ \frac{1+\beta}{1+\alpha} & 1\end{bmatrix}$$

[1] We must be careful at this point with the notation. This is not the scalar standard Jacobi polynomial usually denoted by the same symbol. In fact, if $\mathcal{P}^{(\alpha,\beta)}(z)$ denotes the standard Jacobi polynomials, then $p^{(\alpha,\beta)}(x)=\frac{2^n}{S_n(\alpha,\beta)}\mathcal{P}^{(\beta,\alpha)}(2x-1)$, notice the interchange between the parameters $\alpha\rightleftharpoons\beta$ and the linear transformation of the independent variable x.

we get $F_1 M = -a\tilde{M}$. We finally obtain the monic matrix orthogonal polynomials

$$\tilde{\tilde{P}}_n^{(\alpha,\beta)}(x) = \tilde{M}\begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}\tilde{M}^{-1},$$

for the matrix of measures $\tilde{W}(x)\,\mathrm{d}\mu(x)$ in Example 5.1 of [81] which are explicitly expressed in terms of scalar Jacobi polynomials as follows

$$\tilde{\tilde{P}}_n^{(\alpha,\beta)}(x) =$$
$$\frac{1}{2+\alpha+\beta}\begin{bmatrix} (\alpha+1)p_n^{(\alpha+1,\beta)}(x)+(\beta+1)p_n^{(\alpha,\beta+1)}(x) & -(\alpha+1)(p_n^{(\alpha+1,\beta)}(x)-p_n^{(\alpha,\beta+1)}(x)) \\ -(\beta+1)(p_n^{(\alpha+1,\beta)}(x)-p_n^{(\alpha,\beta+1)}(x)) & (\beta+1)p_n^{(\alpha+1,\beta)}(x)+(\alpha+1)p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}.$$

To conclude with this example let us mention that in [31] it was found that these matrix orthogonal polynomials also satisfy a first order linear ordinary differential equation with matrix polynomials as coefficients. From our point of view this is just a consequence of a remarkable fact regarding the Darboux transformations $p_n^{(\alpha+1,\beta)}(x), p_n^{(\alpha,\beta+1)}(x)$ of the original alternative Jacobi polynomials. Under the hypergeometric function description of the Jacobi polynomials one deduces recurrences for the Jacobi polynomials. In particular, from the Gauss' contiguous relations the first order differential relations appear

$$\begin{aligned}\Big(x\frac{\mathrm{d}}{\mathrm{d}x}+\alpha+1\Big)p_n^{(\alpha+1,\beta)}(x) &= (\alpha+1+n)p_n^{(\alpha,\beta+1)}(x),\\ \Big((x-1)\frac{\mathrm{d}}{\mathrm{d}x}+\beta+1\Big)p_n^{(\alpha,\beta+1)}(x) &= (\beta+1+n)p_n^{(\alpha+1,\beta)}(x).\end{aligned}$$

They can be re-casted as a matrix linear differential equation as follows

$$\begin{multlined}\left(\begin{bmatrix} 0 & x-1 \\ x & 0\end{bmatrix}\frac{\mathrm{d}}{\mathrm{d}x}+\begin{bmatrix} a_1 & \beta+1 \\ \alpha+1 & a_2\end{bmatrix}\right)\begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}\\ = \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}\begin{bmatrix} a_1 & n+\beta+1 \\ n+\alpha+1 & a_2\end{bmatrix},\end{multlined}$$

where $a_1, a_2 \in \mathbb{R}$. This equation is invariant under multiplication on the right and on the left hand side by arbitrary diagonal matrices

$$\begin{multlined}\begin{bmatrix} l_1 & 0 \\ 0 & l_2\end{bmatrix}\left(\begin{bmatrix} 0 & x-1 \\ x & 0\end{bmatrix}\frac{\mathrm{d}}{\mathrm{d}x}+\begin{bmatrix} a_1 & \beta+1 \\ \alpha+1 & a_2\end{bmatrix}\right)\begin{bmatrix} r_1 & 0 \\ 0 & r_2\end{bmatrix}\begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}\\ = \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}\begin{bmatrix} l_1 & 0 \\ 0 & l_2\end{bmatrix}\begin{bmatrix} a_1 & n+\beta+1 \\ n+\alpha+1 & a_2\end{bmatrix}\begin{bmatrix} r_1 & 0 \\ 0 & r_2\end{bmatrix}.\end{multlined}$$

After the similarity transformation, $A \mapsto \tilde{M}^{-\top} A \tilde{M}^{\top}$ we find out that the orthogonal polynomial $\tilde{\hat{P}}_n$ satisfies

$$\Big(A_1(x)\frac{\mathrm{d}}{\mathrm{d}x}+A_0\Big)\big(\tilde{\hat{P}}_n\big)^{\top}=\big(\tilde{\hat{P}}_n\big)^{\top}\Lambda_n,$$

where

$$A_1(x)=\frac{\alpha+1}{\alpha+\beta+2}\left(\begin{bmatrix}-\delta & \delta_-\\ \delta_+ & \delta\end{bmatrix}x+d\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}\right),\quad A_0=\frac{\alpha+1}{\alpha+\beta+2}\begin{bmatrix}C_+-\Delta & \frac{\beta+1}{\alpha+1}(C+\Delta_-)\\ C+\Delta_+ & C_-+\Delta\end{bmatrix},$$

$$\Lambda_n=\frac{\alpha+1}{\alpha+\beta+2}\begin{bmatrix}-\delta & \delta_-\\ \delta_+ & \delta\end{bmatrix}n+A_0,$$

with $d=l_1r_2$ and

$$\begin{aligned}
\delta&=l_1r_2+\frac{\beta+1}{\alpha+1}l_2r_1, & \delta_-&=-l_1r_2+\Big(\frac{\beta+1}{\alpha+1}\Big)^2 l_2r_1, & \delta_+&=l_1r_2-l_2r_1,\\
\Delta&=(\beta+1)(l_1r_2+l_2r_1), & \Delta_-&=-l_1r_2(\alpha+1)+l_2r_1(\beta+1), & \Delta_+&=l_1r_2(\beta+1)-l_2r_1(\alpha+1),\\
C&=-l_1r_1a_1+l_2r_2a_2, & C_-&=\frac{\beta+1}{\alpha+1}l_1r_1a_1+l_2r_2a_2, & C_+&=l_1r_1a_1+\frac{\beta+1}{\alpha+1}l_2r_2a_2.
\end{aligned}$$

If we take $l_1=l_2=-1$, $r_1=r_2=1$, $a_1=\beta+1$ and $a_2=\alpha+1$, we get the first order ordinary differential system in §4 of [31].

Remark 4.15. *The discussion in this example, dealing with the Jacobi polynomials $p_n^{(\alpha+1,\beta)}(x)$ and $p_n^{(\alpha,\beta+1)}(x)$ and the use of the Gauss' contiguous relations, connects with the results in [94], Remark 2.8., see also [92, 93].*

4.16 Example . Here we analyze the Chebyshev case taken from [31], that gives an example of a family of matrix orthogonal polynomials which satisfy a first order linear ordinary differential equation. In §3 of [31] we find a set of monic orthogonal polynomials related to the measure $\tilde{W}(x)\,\mathrm{d}\mu(x)$, where

$$\tilde{W}(x):=\begin{bmatrix}1 & x\\ x & 1\end{bmatrix}, \qquad \mathrm{d}\mu(x)=\frac{1}{\sqrt{1-x^2}}.$$

We have a nonsingular leading coefficient $\left[\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right]$ such that

$$\tilde{W}(x)=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}W(x), \qquad W(x):=I_2x-\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}.$$

Following Remark 4.1 we shall analyze the Darboux transformations $\mathrm{d}\mu(x)\mapsto W(x)\,\mathrm{d}\mu(x)$. Thus, using (4.13) we can write the perturbed monic matrix orthogonal polynomials as follows

$$\hat{P}_n(x)=Q\begin{bmatrix}\frac{t_{n+1}(x)-\frac{t_{n+1}(-1)}{t_n(-1)}t_n(x)}{x+1} & 0\\ 0 & \frac{t_{n+1}(x)-\frac{t_{n+1}(1)}{t_n(1)}t_n(x)}{x-1}\end{bmatrix}Q^{\top}, \qquad Q:=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},$$

where $(t_n(x))_{n\in\mathbb{N}}$ are the monic Chebyshev polynomials of first kind, i.e., $t_n(x)=2^{-n+1}T_n(x)$ with $T_n(x)$ the first kind Chebyshev polynomial of degree n. Therefore, recalling that $T_n(\pm 1)=(\pm 1)^n$ we get

$$\frac{t_{n+1}(\mp 1)}{t_n(\mp 1)}=\mp\frac{1}{2}, \qquad t_{n+1}(x)-\frac{t_{n+1}(\mp 1)}{t_n(\mp 1)}t_n(x)=\frac{1}{2^n}(T_{n+1}(x)\pm T_n(x)).$$

Now, taking into account the mutual recurrence relation satisfied by Chebyshev polynomials of the first and second kind, denoted these last ones by $U_n(x)$, as usual,

$$T_{n+1}(x)=xT_n(x)-(1-x^2)U_{n-1}(x), \qquad T_n(x)=U_n(x)-xU_{n-1}(x),$$

which implies $T_{n+1}(x)=xU_n(x)-U_{n-1}(x)$, we deduce

$$T_{n+1}(x)\pm T_n(x)=(x\pm 1)(U_n(x)\mp U_{n-1}(x)).$$

Consequently,

$$\hat{P}_n(x)=\frac{1}{2^n}Q\begin{bmatrix}U_n(x)-U_{n-1}(x) & 0\\ 0 & U_n(x)+U_{n-1}(x)\end{bmatrix}Q^\top, \qquad Q:=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}.$$

The matrix orthogonal polynomials associated with the original measure $\tilde{W}(x)\,\mathrm{d}\mu(x)$ can be recovered from this by a similarity transformation with $\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$ so that

$$\begin{aligned}\tilde{\hat{P}}_n(x)&=\frac{1}{2^{n+1}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}U_n(x)-U_{n-1}(x) & 0\\ 0 & U_n(x)+U_{n-1}(x)\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\\ &=\frac{1}{2^n}\begin{bmatrix}U_n(x) & -U_{n-1}(x)\\ -U_{n-1}(x) & U_n(x)\end{bmatrix}.\end{aligned}$$

Remark 4.17. *The polynomials $(U_n(x)\mp U_{n-1}(x))_{n\in\mathbb{N}}$ with $U_{-1}(x)=0$, which are orthogonal with respect to the measures $\frac{x\pm 1}{\sqrt{1-x^2}}$, are the well known Chebyshev polynomials of the third and fourth kind, respectively.*

Remark 4.18. *The symmetric structure of the monic orthogonal polynomials can be encoded in the equation*

$$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\tilde{\hat{P}}_n(x)=\tilde{\hat{P}}_n(x)\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}.$$

As in the Jacobi case, the new two scalar families of orthogonal polynomials are related through

$$\Big((x\mp 1)\frac{\mathrm{d}}{\mathrm{d}x}+\frac{1}{2}\Big)\big(U_n(x)\pm U_{n-1}(x)\big)=(n+\frac{1}{2})\big(U_n(x)\mp U_{n-1}(x)\big). \tag{4.15}$$

This follows from

$$(x\mp 1)(U_n'(x)\pm U_{n-1}'(x))=(x\mp 1)\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{T_{n+1}(x)\mp T_n(x)}{x\mp 1}\Big)$$

$$
\begin{aligned}
&= T'_{n+1}(x) \mp T'(x) - \frac{T_{n+1}(x) \mp T_n(x)}{x \mp 1}\\
&= (n+1)U_n(x) - nU_{n-1}(x) - U_n(x) \mp U_{n-1}(x).
\end{aligned}
$$

Here we have used that $T'_n(x) = nU_{n-1}(x)$. The differential equation (4.15) can be written in matrix form

$$
\begin{aligned}
&\left(\begin{bmatrix} 0 & x-1 \\ x+1 & 0 \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}x} + \begin{bmatrix} a_1 & \frac{1}{2} \\ \frac{1}{2} & a_2 \end{bmatrix}\right) \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix}\\
&\qquad = \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix} \begin{bmatrix} a_1 & n+\frac{1}{2} \\ n+\frac{1}{2} & a_2 \end{bmatrix}, \qquad (4.16)
\end{aligned}
$$

where $a_1, a_2 \in \mathbb{R}$ are arbitrary constants. Notice that this matrix equation is invariant under multiplication on the right and on the left hand sides by arbitrary diagonal matrices $L = \operatorname{diag}(l_1, l_2)$ and $R = \operatorname{diag}(r_1, r_2)$,

$$
\begin{aligned}
&\begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \left(\begin{bmatrix} 0 & x-1 \\ x+1 & 0 \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}x} + \begin{bmatrix} a_1 & \frac{1}{2} \\ \frac{1}{2} & a_2 \end{bmatrix}\right) \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix}\\
&\qquad = \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix} \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} a & n+\frac{1}{2} \\ n+\frac{1}{2} & b \end{bmatrix} \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}.
\end{aligned}
$$

After the similarity transformation we find out that the orthogonal polynomial $\tilde{\hat{P}}_n$ satisfies

$$
\left(A_1(x) \frac{\mathrm{d}}{\mathrm{d}x} + A_0\right) \tilde{\hat{P}}_n(x) = \tilde{\hat{P}}_n(x) \Lambda_n,
$$

where

$$
A_1(x) = \begin{bmatrix} -\delta_+ x + \delta_- & \delta_- x - \delta_+ \\ -\delta_- x + \delta_+ & \delta_+ x - \delta_- \end{bmatrix}, \qquad \Lambda_n = \begin{bmatrix} A_+ - \delta_+ \left(n + \frac{1}{2}\right) & A_- + \delta_- \left(n + \frac{1}{2}\right) \\ A_- - \delta_- \left(n + \frac{1}{2}\right) & A_+ + \delta_+ \left(n + \frac{1}{2}\right) \end{bmatrix},
$$

and $A_0 = \Lambda_0$ with

$$
\delta_\pm = l_1 r_2 \pm l_2 r_1, \qquad A_\pm = l_1 r_1 a_1 \pm l_2 r_2 a_2.
$$

Equations (3.1) and (3.2) of §3 of [31] can be recovered choosing

$$
\left(\delta_+ = A_+ = 0, \quad \delta_- = 1, \quad A_- = -\frac{1}{2}\right) \quad \text{and} \quad \left(\delta_- = A_- = 0, \quad \delta_+ = -1, \quad A_+ = \frac{1}{2}\right),
$$

respectively. However, they are all equivalent to (4.16), another form of writing (4.15). This last equation (4.15) is a quite interesting one. Indeed, we have two families of Darboux transformed orthogonal polynomials connected by two first order differential equations. Moreover, we conclude

$$
\left((x^2-1)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (2x \mp 1)\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{4}\right)\big(U_n(x) \mp U_{n-1}(x)\big) = \left(n + \frac{1}{2}\right)^2 \big(U_n(x) \mp U_{n-1}(x)\big).
$$

4.19 Example . *Here we introduce a comment on the matrix Gegenbauer matrix valued polynomials discussed in [94]. In this case the matrix of weights is a symmetric matrix, $W^{(\nu)}:[-1,1]\to\mathbb{R}^{N\times N}$, with matrix coefficients*

$$(W^{(\nu)}(x))_{i,j} := (1-x^2)^{\nu-1/2} \sum_{k=\max(0,i+j+1-N)} \alpha_k^{(\nu)}(i,j) C_{i+j-2k}^{(\nu)}(x), \qquad i\geq j,$$

where $\alpha_t^{(\nu)}(i,j)$ are real numbers and $C_n^{(\nu)}(x)$ stands for the Gegenbauer or ultraspherical polynomials. E. Koelink and P. Román kindly communicated us a nice feature of the matrix Christoffel transformation discussed in this dissertation when acting on this family of monic orthogonal polynomial: two families of Gegenbauer monic orthogonal polynomial associated with matrices of weights $W^{(\nu_1)}(x)$ and $W^{(\nu_2)}(x)$, such that $\nu_1-\nu_2=m\in\mathbb{Z}$, are linked by a matrix Christoffel transformation. Now, the perturbing polynomial $W(x)$ has $\deg W=2m$. These examples are, in general, reducible to two irreducible blocks of sizes $N/2$, for N even, and $(N+1)/2$ and $(N-1)/2$ for odd N respectively. For a discussion on the orthogonal and non orthogonal reducibility of these examples see [94, 95].

4.20 Example . *Let $(L_n^\alpha(x))_{n\in\mathbb{N}}$ be the sequence of monic Laguerre polynomials which are orthogonal with respect to the measure $d\mu=e^{-x}x^\alpha dx$, $\alpha>-1$, supported on $(0,\infty)$ and $K_n^\alpha(x,y)$ the corresponding kernel polynomial of degree n with parameter α. These polynomials satisfy the following properties*

Proposition 4.21. *Let $(L_n^\alpha(x))_{n\in\mathbb{N}}$ and $(L_n^\beta(x))_{n\in\mathbb{N}}$ be the monic Laguerre polynomials of parameter α and β, respectively, and let $(a)_k=a(a+1)+\dots(a+k-1)$ with $a\in\mathbb{C}$, $k\geq 1$, and $(a)_0=1$ be the Pochhammer symbol. Then, for $n\in\mathbb{N}$ we have*

i) $L_n^\alpha(x)=\sum_{k=0}^n(-1)^{n-k}\binom{n}{k}(\alpha-\beta)_{n-k}L_k^\beta(x)$ *(Connection formula).*

ii) $(L_n^\alpha)^{(i)}(0)=(-1)^{n+i}\frac{n!\Gamma(\alpha+n+1)}{(n-i)!\Gamma(\alpha+i+1)}$.

iii) $\int_0^\infty L_n^\alpha(x)L_m^\alpha(x)d\mu=n!\Gamma(n+\alpha+1)\delta_{n,m}$.

iv) $L_n^{\alpha+1}(x)=\frac{\|L_n^\alpha(x)\|_\alpha^2}{L_n^\alpha(0)}K_n^\alpha(x,0)$.

v) $xL_n^{\alpha+1}(x)=L_{n+1}^\alpha(x)+(n+\alpha+1)L_n^\alpha(x)$.

vi) $nL_{n-1}^{\alpha+1}(x)=L_n^\alpha(x)-L_n^{\alpha+1}(x)$

vii) $K_n^\alpha(x,y)=yK_{n-1}^{\alpha+1}(x,y)+\frac{L_n^{\alpha+1}(x)L_n^\alpha(y)}{\|L_n^\alpha(x)\|_\alpha^2}$.

Proof. We only prove vii), the other are well known in the literature (see [140]). Observe that

$$yK_{n-1}^{\alpha+1}(x,y)+\frac{L_n^{\alpha+1}(x)L_n^\alpha(y)}{\|L_n^\alpha(x)\|_\alpha^2}=\frac{nyL_n^{\alpha+1}(x)L_{n-1}^{\alpha+1}(y)-nyL_n^{\alpha+1}(y)L_{n-1}^{\alpha+1}(x)+(x-y)L_n^{\alpha+1}(x)L_n^\alpha(y)}{(x-y)\|L_n^\alpha\|_\alpha^2}$$

Using vi) we have

$$nyL_n^{\alpha+1}(x)L_{n-1}^{\alpha+1}(y)-nyL_n^{\alpha+1}(y)L_{n-1}^{\alpha+1}(x)=yL_n^{\alpha+1}(x)L_n^{\alpha+1}(y)-yL_n^{\alpha+1}(y)L_n^{\alpha+1}(x)$$

Thus

$$yK_{n-1}^{\alpha+1}(x,y)+\frac{L_n^{\alpha+1}(x)L_n^{\alpha}(y)}{\|L_n^{\alpha}(x)\|_{\alpha}^2}=\frac{xL_n^{\alpha+1}(x)L_n^{\alpha}(y)-yL_n^{\alpha+1}(y)L_n^{\alpha}(x)}{(x-y)\|L_n^{\alpha}\|^2}$$

and using v) we get the result. ■

If we take $W(x)=xI_2-\begin{bmatrix}0&1\\0&0\end{bmatrix}$*, then using Proposition 4.21 and (4.14) we get that the sequences of monic matrix bi-orthogonal polynomials* $(\hat{\mathcal{L}}_n^{\alpha[1]}(x),\hat{\mathcal{L}}_n^{\alpha[2]}(x))_{n\in\mathbb{N}}$ *with respect to the new matrix measure* $d\hat{\mu}=W(x)e^{-x}x^{\alpha}dx$ *are*

$$\hat{\mathcal{L}}_n^{\alpha[1]}(x)=\begin{bmatrix}L_n^{\alpha+1}(x) & \frac{L_n^{\alpha+1}(x)-\frac{(\alpha+n+1)}{(\alpha+1)}L_n^{\alpha}(x)}{x}\\ & L_n^{\alpha+1}(x)\end{bmatrix}=\begin{bmatrix}L_n^{\alpha+1}(x) & -\frac{n}{\alpha+1}L_{n-1}^{\alpha+2}(x)\\ & L_n^{\alpha+1}(x)\end{bmatrix},$$

$$\hat{\mathcal{L}}_n^{\alpha[2]}(x)(\hat{H}_n)^{-1}=\begin{bmatrix}\frac{L_n^{(\alpha+1)}(x)}{n!\Gamma(\alpha+n+1)} & \frac{(-1)^{n+1}L_n^{\alpha+1}(x)}{(n-1)!\Gamma(\alpha+1)}-\frac{L_n^{\alpha+1}(x)}{xn!\Gamma(n+\alpha+1)}+\frac{nL_{n+1}^{\alpha}(x)}{x(\alpha+1)}-\frac{(n+1)(\alpha+n+1)L_n^{\alpha}(x)}{x(\alpha+1)}\\ & \frac{L_n^{(\alpha+1)}(x)}{n!\Gamma(\alpha+n+1)}\end{bmatrix}.$$

4.3 Singular leading coefficient matrix polynomial perturbations

After studying some examples that the literature provides us one may realize that, even thought it is generic to assume the perturbing matrix polynomial $W(x)$ to have a nonsingular leading coefficient, many examples have a singular matrix as its leading coefficient. This situation is a special feature of the matrix case setting since in the scalar case, having a singular leading term would mean that this coefficient is just zero (affecting, of course, to the degree of the polynomial). For this reason, when dealing with this kind of matrix polynomials their degree should make no sense. The effect that this fact has on our reasoning is that since $\deg[\det W(x)]\leq Np$ the information encoded in the zeros (and the corresponding adapted polynomials) of $\det W(x)$ is no longer enough to make the matrices of Theorem 4.11 of the needed size. Therefore, there will be no way to express the perturbed polynomials just in terms of the initial ones evaluated at the zeros of $\det W(x)$ and the method to find a Christoffel type formula fails. However, the information that seems to be missing in these cases may actually not be necessary due to the singular character of the leading coefficient of the perturbing polynomial. Let us consider the following example to take a glimpse of this scenario.

Let us pick up a scalar measure $\mathrm{d}\mu(x)$ and its associated monic orthogonal polynomials $\big(p_k(x)\big)_{k\in\mathbb{N}}$ together with their norms and three term recurrence relation

$$h_k\delta_{kj} := \langle p_k, p_j\rangle, \qquad xp_k(x) = J_{k,k-1}p_{k-1}(x) + J_{k,k}p_k(x) + p_{k+1}(x), \qquad J_{k,k-1} = \frac{h_k}{h_{k-1}} > 0.$$

Now, consider its $2q\times 2q$ matrix diagonal extension in $\mathbb{R}^{2q\times 2q}[x]$

$$P_k(x) := p_k(x)I_{2q}, \qquad\qquad H_k := h_k I_{2q}.$$

Our aim is to consider the following matrix polynomial (with singular leading coefficient)

$$W(x) := \begin{bmatrix} I_q + AA^\top x^2 & Ax \\ A^\top x & I_q \end{bmatrix}, \qquad\qquad A\in\mathbb{R}^{q\times q},$$

which is inspired by the $q=1$ case $\left[\begin{smallmatrix} 1+a^2x^2 & ax \\ ax & 1\end{smallmatrix}\right]$ (see [58], and references therein). We will study the corresponding perturbations of our initial scalar measure i.e., $\mathrm{d}\hat{\mu}(x) := W(x)\,\mathrm{d}\mu(x)$ in order to obtain the transformed matrix orthogonal polynomials

$$\hat{P}_k(x) := \begin{bmatrix} (\hat{P}_k)_{1,1} & (\hat{P}_k)_{1,2} \\ (\hat{P}_k)_{2,1} & (\hat{P}_k)_{2,2}\end{bmatrix}, \qquad \hat{P}_k(x)\in\mathbb{R}^{2q\times 2q}[x], (\hat{P}_k)_{i,j}\in\mathbb{R}^{q\times q}[x],$$

$$\langle \hat{P}_k, \hat{P}_j\rangle_{\hat{\mu}} := \delta_{k,j}\hat{H}_k = \delta_{k,j}\begin{bmatrix} (\hat{H}_k)_{1,1} & (\hat{H}_k)_{1,2} \\ (\hat{H}_k)_{2,1} & (\hat{H}_k)_{2,2}\end{bmatrix}, \qquad \hat{H}_k\in\mathbb{R}^{2q\times 2q}, (\hat{H}_k)_{i,j}\in\mathbb{R}^{q\times q}.$$

We have splitted them in this way for computational purposes. Notice that since $W(x) = W(x)^\top$ we have $\hat{M} = \hat{M}^\top := \hat{S}^{-1}\hat{H}[\hat{S}^{-1}]^\top$ and, therefore, $\hat{P}^{[1]} = \hat{P}^{[2]} := \hat{P}$ and $\hat{H}_k = (\hat{H}_k)^\top$.
Let us point out that

$$W(x) = \mathcal{W}(x)\mathcal{W}(x)^\top, \qquad \mathcal{W} := \begin{bmatrix} I_q & Ax \\ 0 & I_q\end{bmatrix}, \qquad \mathcal{W}^{-1} = \begin{bmatrix} I_q & -Ax \\ 0 & I_q\end{bmatrix}.$$

This implies that $\det W = \det\mathcal{W} = 1$ and, consequently, there is no spectral analysis to perform as there are non eigenvalues at all. Thus, the relation between the original and perturbed measures and moment matrices is

$$[\mathcal{W}(x)]^{-1}\,\mathrm{d}\hat{\mu} = \mathrm{d}\mu[\mathcal{W}(x)]^\top, \qquad\qquad [\mathcal{W}(\Lambda)]^{-1}\hat{M} = M[\mathcal{W}(\Lambda)]^\top.$$

Definition 4.22. *We introduce the resolvent or connection matrix*

$$\omega := \hat{S}\mathcal{W}(\Lambda)S^{-1}.$$

Proposition 4.23. *The matrix ω is block tridiagonal, having nonzeros entries only in its diagonal and first superdiagonal and subdiagonal, and satisfies*

$$\omega^{-1} = H\omega^\top\hat{H}^{-1}.$$

Moreover, we have the connection formula

$$\omega P = \hat{P}\mathcal{W}(x).$$

Proof. The first relation is a consequence of the LU factorization of the moment matrices and the connection formula is a straightforward consequence of the definition of ω. ■

Proposition 4.24. i) *The matrices*

$$\rho_{k+1} := \left(I_q + J_{k+1,k}A^\top A\right)^{-1}, \qquad k \in \{-1,0,1,\dots\},$$

exist.

ii) *The perturbed monic orthogonal polynomial can be written in terms of the original ones as follows*

$$\hat{P}_{k+1}(x)\mathcal{W}(x) = \\ -\begin{bmatrix} J_{k+1,k}J_{k+1,k+1}A\rho_{k+1}A^\top & 0 \\ J_{k+1,k}\rho_{k+1}A^\top & 0 \end{bmatrix} p_k(x) + \begin{bmatrix} I_q & J_{k+1,k+1}A\rho_{k+1} \\ 0 & \rho_{k+1} \end{bmatrix} p_{k+1}(x) + \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix} p_{k+2}(x),$$

for $k \in \{-1,0,1,\dots\}$.

Proof. From the $(k+1)$-th row of the connection formula we have that

$$\omega_{k+1,k}p_k(x) + \omega_{k+1,k+1}p_{k+1}(x) + \omega_{k+1,k+2}p_{k+2}(x) = \hat{P}_{k+1}(x)\mathcal{W}(x),$$

but from Definition 4.22 and Proposition 4.23 one realizes that the previous expression reads

$$\hat{P}_{k+1}(x)\mathcal{W}(x) = \hat{H}_{k+1}\begin{bmatrix} 0 & -A \\ 0 & 0 \end{bmatrix}^\top h_k^{-1}p_k(x) + \begin{bmatrix} (\omega_{k+1,k+2})_{11} & (\omega_{k+1,k+2})_{12} \\ (\omega_{k+1,k+2})_{21} & (\omega_{k+1,k+2})_{22} \end{bmatrix} p_{k+1}(x) + \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix} p_{k+2}(x).$$

Now, taking into account that both $(\hat{P}_{k+1})_{11}, (\hat{P}_{k+1})_{22}$ are monic $q \times q$ polynomials of degree $k+1$, while $(\hat{P}_{k+1})_{12}, (\hat{P}_{k+1})_{21}$ are $q \times q$ polynomials of degree less than $k+1$, it is not so hard to see after using the three term recurrence relation of the initial polynomials that

$$\begin{aligned} (\omega_{k+1,k+2})_{11} &= I_q, & (\omega_{k+1,k+2})_{12} &= J_{k+1,k+1}A - h_k^{-1}(\hat{H}_{k+1})_{12}A^\top A, \\ (\omega_{k+1,k+2})_{21} &= 0, & (\omega_{k+1,k+2})_{22} &= I_q - h_k^{-1}(\hat{H}_{k+1})_{22}A^\top A. \end{aligned}$$

Hence, every coefficient that appears in the connection formula in terms of the still unknown norms of the monic orthogonal polynomials is given. Therefore, we just need to compute the second block column of the following integral

$$\begin{aligned} \int [\omega_{k+1,k}p_k(x) + \omega_{k+1,k+1}p_{k+1}(x) + \omega_{k+1,k+2}p_{k+2}(x)] \left[(\mathcal{W}(x))^\top x^{k+1}\right] \mathrm{d}\mu(x) \\ = \int \hat{P}_{k+1}(x)\mathcal{W}(x)\left[(\mathcal{W}(x))^\top x^{k+1}\right] \mathrm{d}\mu(x) \\ = \int \hat{P}_{k+1}(x)\,\mathrm{d}\hat{\mu}(x)x^{k+1} \\ = \hat{H}_{k+1}, \end{aligned}$$

which yields

$$(\hat{H}_{k+1})_{12} = J_{k+1,k+1}h_{k+1}A\left(I_q + J_{k+1,k}A^\top A\right)^{-1}, \qquad (\hat{H}_{k+1})_{22} = h_{k+1}\left(I_q + J_{k+1,k}A^\top A\right)^{-1}.$$

Therefore,

$$(\omega_{k+1,k+2})_{12} = J_{k+1,k+1}A\left(I_q + J_{k+1,k}A^\top A\right)^{-1}, \qquad (\omega_{k+1,k+2})_{22} = \left(I_q + J_{k+1,k}A^\top A\right)^{-1},$$

and the stated result follows. ■

For $q = 1$ and the classical measures we have, see [62],

Corollary 4.25. i) *Hermite monic polynomials* $\left(\mathcal{H}_k(x)\right)_{k\in\mathbb{N}}$ *with norm* $h_k = \pi^{\frac{1}{2}}\frac{k!}{2^k}$

$$J_{k+1,k} = \frac{k+1}{2}, \qquad J_{k+1,k+1} = 0, \qquad \rho_{k+1} := \frac{2}{2+a^2(k+1)}.$$

ii) *Laguerre monic polynomials* $\left(\mathcal{L}_k^\alpha(x)\right)_{k\in\mathbb{N}}$ *with norm* $h_k = k!\Gamma(k+1+\alpha)$

$$J_{k+1,k} = (k+1)(k+\alpha+2), \quad J_{k+1,k+1} = (2k+\alpha+3), \quad \rho_{k+1} := \frac{1}{1+a^2(k+1)(k+1+\alpha)}.$$

iii) *Jacobi monic polynomials* $\left(\mathcal{P}_k^{(\alpha,\beta)}(x)\right)_{k\in\mathbb{N}}$ *with norm* $h_k = \frac{2^{2k+\alpha+\beta+1}k!\Gamma(k+\alpha+\beta+1)\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{\Gamma(2k+\alpha+\beta+1)^2(2k+\alpha+\beta+1)}$

$$\begin{aligned}
J_{k+1,k} &= \frac{4(k+1)(k+\alpha+1)(k+\beta+1)(k+\alpha+\beta+1)}{(2k+\alpha+\beta+2)^2(2k+\alpha+\beta+3)(2k+\alpha+\beta+1)},\\
J_{k+1,k+1} &= \frac{\beta^2-\alpha^2}{(2k+\alpha+\beta+2)(2k+\alpha+\beta+4)},\\
\rho_{k+1} &:= \left(1+a^2\frac{4(k+1)(k+\alpha+1)(k+\beta+1)(k+\alpha+\beta+1)}{(2k+\alpha+\beta+2)^2(2k+\alpha+\beta+3)(2k+\alpha+\beta+1)}\right)^{-1}.
\end{aligned}$$

Perturbing them by the matrix polynomial

$$W(x) = \mathcal{W}(x)\left(\mathcal{W}(x)\right)^\top, \qquad \mathcal{W}(x) = \begin{bmatrix} 1 & ax \\ 0 & 1 \end{bmatrix},$$

one obtains the perturbed monic orthogonal polynomials related to the classical orthogonal polynomials as follows

$$\begin{aligned}
\hat{\mathcal{H}}_{k+1}(x)\mathcal{W}(x) &= \begin{bmatrix} 0 & 0 \\ \frac{\rho_{k+1}-1}{a} & 0 \end{bmatrix}\mathcal{H}_k(x) + \begin{bmatrix} 1 & 0 \\ 0 & \rho_{k+1} \end{bmatrix}\mathcal{H}_{k+1}(x) + \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}\mathcal{H}_{k+2}(x),\\
\hat{\mathcal{L}}^\alpha_{k+1}(x)\mathcal{W}(x) &= \begin{bmatrix} (\rho_{k+1}-1)(2k+3+\alpha) & 0 \\ \frac{\rho_{k+1}-1}{a} & 0 \end{bmatrix}\mathcal{L}^\alpha_k(x) + \begin{bmatrix} 1 & a(2k+3+\alpha)\rho_{k+1} \\ 0 & \rho_{k+1} \end{bmatrix}\mathcal{L}^\alpha_{k+1}(x)
\end{aligned}$$

$$+\begin{bmatrix}0 & a\\ 0 & 0\end{bmatrix}\mathcal{L}^{\alpha}_{k+2}(x).$$

$$\hat{\mathcal{P}}^{(\alpha,\beta)}_{k+1}(x)\mathcal{W}(x)=-\begin{bmatrix}J_{k+1,k}J_{k+1,k+1}a^2\rho_{k+1} & 0\\ aJ_{k+1,k}\rho_{k+1} & 0\end{bmatrix}\mathcal{P}^{(\alpha,\beta)}_{k}(x)+\begin{bmatrix}1 & aJ_{k+1,k+1}\rho_{k+1}\\ 0 & \rho_{k+1}\end{bmatrix}\mathcal{P}^{(\alpha,\beta)}_{k+1}(x)$$
$$+\begin{bmatrix}0 & a\\ 0 & 0\end{bmatrix}\mathcal{P}^{(\alpha,\beta)}_{k+2}(x).$$

4.4 Extension to non-Abelian 2D Toda hierarchies

Matrix orthogonal polynomials are connected with non-Abelian Toda lattices, see [14, 122].

4.4.1 Block Hankel moment matrices vs multi-component Toda hierarchies

Let us take $M=(m_{i,j})_{i,j=0}^{\infty}$, $m_{i,j}\in\mathbb{R}^{p\times p}$, a semi-infinite block matrix having a Gaussian factorization

$$M=(S_1)^{-1}H(S_2)^{-\top},$$

where S_1,S_2 are lower uni-triangular block matrices and H is block diagonal. Notice that conditions for the existence of this factorization were given in Proposition 1.52.

Definition 4.26. *We introduce some continuous flows or perturbations of this semi-infinite matrix. First, let us consider we first consider the diagonal matrices*

$$t_{i,j}=\operatorname{diag}(t_{i,j,1},\dots,t_{i,j,p})\in\mathbb{R}^{p\times p},\qquad i=1,2,\qquad j\in\mathbb{N}\setminus\{0\},$$

the semi-infinite undressed wave matrices

$$V_i^{(0)}(t_i):=\exp\Big(\sum_{j=0}^{\infty}t_{i,j}\Lambda^j\Big),\qquad i=1,2,$$

and the perturbed matrix $M(t)$, $t=(t_1,t_2)$, $t_i=\{t_{i,j,a}\}_{\substack{j\in\mathbb{N}\setminus\{0\}\\ a\in\{1,\dots,p\}}}$

$$M(t)=V_1^{(0)}(t_1)M\big(V_2^{(0)}(t_2)\big)^{-\top}.$$

Observe that we do not require any Hankel form for the matrix M, modelled by $\Lambda M=M\Lambda^{\top}$. However, if $M(0)$ is a Hankel matrix, then $M(t)$ is also a Hankel matrix taking into account $\Lambda M(t)=M(t)\Lambda^{\top}$. Hence, if $\mathrm{d}\mu(x)$ is the initial matrix of measures, then the new matrix of measures $\mathrm{d}\mu(x,t)$ is

$$\mathrm{d}\mu(x,t)=\exp\Big(\sum_{j=0}^{\infty}t_{1,j}x^j\Big)\,\mathrm{d}\mu(x)\exp\Big(-\sum_{j=0}^{\infty}t_{2,j}x^j\Big).$$

Here $M(t)$ will be the moment matrix of the matrix of measures. Moreover, if at any time the matrix of measures is block Hankel, then it will be a Hankel block matrix at any time. If we assume that the Gaussian factorization exists, again we can write

$$M(t) = (S_1(t))^{-1}H(t)(S_2(t))^{-\top}.$$

As we know, for the block Hankel case we are dealing with bi-orthogonal or orthogonal polynomials with respect to the associated matrix of measures. What happens in the general case? Following [2] and [105] we can understand the Gaussian factorization also as a bi-orthogonality condition. The semi-infinite vectors of polynomials will be

$$P^{[1]}(x) := S_1(t)\chi(x), \qquad\qquad P^{[2]}(x) := S_2(t)\chi(x),$$

and we consider a sesquilinear form in $\mathbb{R}^{p\times p}[x]$, see Section 1.3, such that for any pair of matrix polynomials $P = \sum_{k=0}^{\deg P} p_k x^k$ and $Q(x) = \sum_{l=0}^{\deg Q} q_l x^l$

$$\langle P(x), Q(x)\rangle = \sum_{\substack{k=1,\dots,\deg P\\ l=1,\dots,\deg Q}} p_k M_{k,l}(t)(q_l)^\top,$$

where

$$M_{k,l}(t) = \left\langle x^k I_p, x^l I_p \right\rangle$$

can be read as the Gram matrix of the sesquilinear form. With respect to this sesquilinear form we have the bi-orthogonality condition

$$\left\langle P_k^{[1]}(x), P_l^{[2]}(x) \right\rangle = H_k(t)\delta_{k,l}.$$

For a block Hankel initial condition this sesquilinear form is associated with a linear functional of a matrix of measures. In [13] different examples are discussed in the matrix orthogonal polynomials scenario. For example, multigraded Hankel matrices M fulfilling

$$\Big(\sum_{a=1}^{p} \Lambda^{n_a} E_{a,a}\Big) M = M \Big(\sum_{a=1}^{p} \left(\Lambda^\top\right)^{m_a} E_{a,a}\Big),$$

where $n_1,\dots,n_p, m_1,\dots,m_p$, are positive integers and $E_{i,j}$ is the matrix with 1 in the (i,j) entry and zero elsewhere, can be realized as

$$M_{k,l} = \int x^k \,\mathrm{d}\mu^{(l)}(x)$$

in terms of matrices of measures $\mathrm{d}\mu^{(l)}(x)$ which satisfy the following periodicity condition

$$\mathrm{d}\mu_{a,b}^{(l+m_a)}(x) = x^{n_a}\,\mathrm{d}\mu_{a,b}^{(l)}(x). \tag{4.17}$$

Therefore, given the measures $\mathrm{d}\mu^{(0)}_{a,b},\dots,\mathrm{d}\mu^{(m_b-1)}_{a,b}$ we can recover all the others from (4.17). In this case, we have generalized orthogonality conditions like

$$\int P^{[1]}_k(x)\,\mathrm{d}\mu^{(l)}(x)=0, \qquad l=0,\dots,k-1.$$

Coming back to the Gaussian factorization, we consider the wave matrices

$$\begin{aligned}V_1(t):=&S_1(t)V^{(0)}_1(t_1),\\ \tilde V_2(t):=&\tilde S_2(t)(V^{(0)}_2(t_2))^\top,\end{aligned}$$

where $\tilde S_2(t):=H(t)(S_2(t))^{-\top}$.

Proposition 4.27. *The wave matrices satisfy*

$$\big(V_1(t)\big)^{-1}\tilde V_2(t)=M(t). \tag{4.18}$$

Proof. It is a straightforward consequence of the Gaussian factorization. ■

Given a semi-infinite matrix A we have an unique splitting $A=A_++A_-$, where A_+ is an upper triangular block matrix and A_- is a strictly lower triangular block matrix.

Proposition 4.28. *The following equations hold*

$$\begin{aligned}&\frac{\partial S_1}{\partial t_{1,j,a}}(S_1)^{-1}=-\Big(S_1E_{a,a}\Lambda^j(S_1)^{-1}\Big)_-, && \frac{\partial S_1}{\partial t_{2,j,a}}(S_1)^{-1}=\Big(\tilde S_2E_{a,a}\big(\Lambda^\top\big)^j(\tilde S_2)^{-1}\Big)_-,\\ &\frac{\partial \tilde S_2}{\partial t_{1,j,a}}(\tilde S_2)^{-1}=\Big(S_1E_{a,a}\Lambda^j(S_1)^{-1}\Big)_+, && \frac{\partial \tilde S_2}{\partial t_{2,j,a}}(\tilde S_2)^{-1}=-\Big(\tilde S_2E_{a,a}\big(\Lambda^\top\big)^j(\tilde S_2)^{-1}\Big)_+.\end{aligned}$$

Proof. Taking right derivatives in (4.18)

$$\frac{\partial V_1}{\partial t_{i,j,a}}(V_1)^{-1}=\frac{\partial \tilde V_2}{\partial t_{i,j,a}}(\tilde V_2)^{-1}, \qquad i\in\{1,2\}, \qquad j\in\mathbb{N}\setminus\{0\},$$

where

$$\begin{aligned}&\frac{\partial V_1}{\partial t_{1,j,a}}(V_1)^{-1}=\frac{\partial S_1}{\partial t_{1,j,a}}(S_1)^{-1}+S_1E_{a,a}\Lambda^j(S_1)^{-1}, && \frac{\partial V_1}{\partial t_{2,j,a}}(V_1)^{-1}=\frac{\partial S_1}{\partial t_{2,j,a}}(S_1)^{-1},\\ &\frac{\partial \tilde V_2}{\partial t_{1,j,a}}(\tilde V_2)^{-1}=\frac{\partial \tilde S_2}{\partial t_{1,j,a}}(\tilde S_2)^{-1}, && \frac{\partial \tilde V_2}{\partial t_{2,j,a}}(\tilde V_2)^{-1}=\frac{\partial \tilde S_2}{\partial t_{2,j,a}}(\tilde S_2)^{-1}+\tilde S_2E_{a,a}\big(\Lambda^\top\big)^j(\tilde S_2)^{-1},\end{aligned}$$

and the result follows immediately. ■

As a consequence,

Proposition 4.29. *The multicomponent 2D Toda lattice equations*

$$\frac{\partial}{\partial t_{2,1,b}}\Big(\frac{\partial H_k}{\partial t_{1,1,a}}(H_k)^{-1}\Big)+E_{a,a}H_{k+1}E_{b,b}(H_k)^{-1}-H_kE_{b,b}(H_{k-1})^{-1}E_{a,a}=0.$$

hold.

Proof. From Proposition 4.28 we get

$$\frac{\partial H_k}{\partial t_{1,1,a}}(H_k)^{-1}=\beta_kE_{a,a}-E_{a,a}\beta_{k+1}, \qquad \frac{\partial \beta_k}{\partial t_{2,1,b}}=H_kE_{b,b}(H_{k-1})^{-1},$$

where $\beta_k\in\mathbb{R}^{p\times p}$, $k=1,2,\dots$, are the first subdiagonal coefficients in S_1. ■

The multi-component Toda and KP hierarchies were introduced in [142]. In [104, 105] its relevance in integrable aspects of differential geometry was emphasized, and in [88] a representation approach was developed, while in [1, 14] it was used in relation with multiple orthogonality. A comprehensive approach to multi-component 2D Toda hierarchy with applications in dispersionless integrability or generalized orthogonal polynomials can be found in [13, 106, 107].
If we introduce the total flows given by the derivatives

$$\partial_{i,j}:=\sum_{a=1}^{p}\frac{\partial}{\partial t_{i,j,a}}$$

we get the non-Abelian 2D Toda lattice

$$\partial_{2,1}\big(\partial_{1,1}(H_k)\cdot(H_k)^{-1}\big)+H_{k+1}(H_k)^{-1}-H_k(H_{k-1})^{-1}=0.$$

The non-Abelian Toda lattice was introduced in the context of string theory by Polyakov, (see [125, 126]), and then studied under the inverse spectral transform, by Mikhailov [121], and Riemann surface theory, by Krichever [101]. The Darboux transformations were considered in [130] and later on [123].
The non-Abelian 2D Toda lattice hierarchy is a reduction of the multicomponent hierarchy by taking the diagonal time matrices $t_{i,j}=\operatorname{diag}(t_{i,j,1},\dots,t_{i,j,p})$ proportional to the identity, i.e.

$$t_{i,j}\mapsto t_{i,j}I_p, \qquad t_{i,j}\in\mathbb{R}.$$

These equations are just the first members of an infinite set of nonlinear partial differential equations, an integrable hierarchy.

Definition 4.30. *Given two semi-infinite block matrices A,B, $[A,B]=AB-BA$ stands for the usual commutator of matrices.*

Definition 4.31. *The partial, Lax and Zakharov–Shabat matrices are given by*

$$\begin{aligned}\Pi_{1,a}&:=S_1E_{a,a}(S_1)^{-1}, & \Pi_{2,a}&:=\tilde S_2E_{a,a}(\tilde S_2)^{-1},\\ L_1&:=S_1\Lambda(S_1)^{-1}, & L_2&:=\tilde S_2\Lambda^{\top}(\tilde S_2)^{-1},\\ B_{1,j,a}&:=\big(\Pi_{1,a}(L_1)^j\big)_+, & B_{2,j,a}&:=\big(\Pi_{2,a}(L_2)^j\big)_-.\end{aligned}$$

Proposition 4.32 (The integrable hierarchy)**.** *The wave matrices obey the evolutionary linear systems*

$$
\begin{aligned}
\frac{\partial V_1}{\partial t_{1,j,a}} &= B_{1,j,a}V_1, & \frac{\partial V_1}{\partial t_{2,j,a}} &= B_{2,j,a}V_1,\\
\frac{\partial \tilde{V}_2}{\partial t_{1,j,a}} &= B_{1,j,a}\tilde{V}_2, & \frac{\partial \tilde{V}_2}{\partial t_{2,j,a}} &= B_{2,j,a}\tilde{V}_2,
\end{aligned}
$$

the partial and Lax matrices satisfy the following Lax equations

$$
\frac{\partial \Pi_{i',a'}}{\partial t_{i,j,a}} = \Big[B_{i,j,a},\Pi_{i',a'}\Big], \qquad \frac{\partial L_{i'}}{\partial t_{i,j,a}} = \Big[B_{i,j,a},L_{i'}\Big],
$$

and Zakharov–Sabat matrices fulfill the following Zakharov–Shabat equations

$$
\frac{\partial B_{i',j',a'}}{\partial t_{i,j,a}} - \frac{\partial B_{i,j,a}}{\partial t_{i',j',a'}} + \big[B_{i,j,a},B_{i',j',a'}\big] = 0.
$$

Proof. It follows from Proposition 4.28. ∎

A key observation, regarding orthogonal polynomials, must be pointed out. When orthogonal polynomials are involved, and the matrices to factorize are block Hankel, equivalently $\Lambda M = M\Lambda^\top$, we get $L_1 = S_1\Lambda S_1^{-1} = \tilde{S}_2\Lambda^\top\tilde{S}_2^{-1} = L_2$. As the reader may have noticed the Lax matrices L_1 and L_2 are, by construction, lower and upper Hessenberg block matrices, respectively. However, when the Hankel property holds, both Lax matrices are equal

$$
L_1 = L_2
$$

and, therefore, we are dealing with a tridiagonal block matrix, i.e., a Jacobi block matrix. Moreover, this Hankel condition implies an invariance property under the flows introduced above, as we have that $M(t) = V_1^{(0)}(t_1 - t_2)M$, i.e. there are only one type of flows. This condition also implies that for the total flows we have

$$
\begin{aligned}
(\partial_{1,j}+\partial_{2,j})V_1 &= V_1\Lambda^j, & (\partial_{1,j}+\partial_{2,j})\tilde{V}_2 &= \tilde{V}_2(\Lambda^\top)^j,\\
(\partial_{1,j}+\partial_{2,j})L_1 &= 0, & (\partial_{1,j}+\partial_{2,j})L_2 &= 0.
\end{aligned}
$$

Therefore, in the block Hankel case we are dealing with the multicomponent 1D Toda hierarchy.

4.4.2 The Christoffel transformation for the non-Abelian 2D Toda hierarchy

The idea is to follow what we did in §4.1.1 and consider an initial condition $\hat{M}$ at $t = 0$, this is

$$
\hat{M} = W(\Lambda)M
$$

for a matrix polynomial $W(t) \in \mathbb{R}^{p\times p}[t]$. Observe that using the scalar times $t_{i,j} \in \mathbb{R}$ of the non-Abelian flows determined by

$$V_i^{(0)} := \exp\Big(\sum_{j=0}^{\infty} t_{i,j}\Lambda^j\Big), \qquad\qquad i=1,2,$$

the perturbed matrix is

$$\begin{aligned}\hat{M}(t) =& V_1^{(0)}(t_1)\hat{M}\big(V_2^{(0)}(t_2)\big)^{-\top}\\ =& W(\Lambda)M(t).\end{aligned}$$

Here we have used that $[W(\Lambda), V_1^{(0)}(t)] = 0$, $\forall t_{1,j} \in \mathbb{R}$. Let us stress that we could request only $t_{1,j}$ to be scalars and $t_{2,j}$ to be diagonal matrices. Despite this is a more general situation, we prefer to show how the method works in this simpler scenario.
Assuming that the block Gauss factorization holds, we proceed as in §4.1.1 and introduce the resolvents

$$\omega^{[1]}(t) := \hat{S}_1(t)W(\Lambda)\big(S_1(t)\big)^{-1}, \qquad\qquad \omega^{[2]}(t) := \big(S_2(t)\big(\hat{S}_2(t)\big)^{-1}\big)^{\top}.$$

From the LU factorization we get

$$\big(\hat{S}_1(t)\big)^{-1}\hat{H}(t)\big(\hat{S}_2(t)\big)^{-\top} = W(\Lambda)\big(S_1(t)\big)^{-1}H(t)\big(S_2(t)\big)^{-\top},$$

so that

$$\hat{H}(t)\big(S_2(t)\big(\hat{S}_2(t)\big)^{-1}\big)^{\top} = \hat{S}_1(t)W(\Lambda)\big(S_1(t)\big)^{-1}H(t),$$

and, consequently,

$$\hat{H}(t)\omega^{[2]}(t) = \omega^{[1]}(t)H(t)$$

holds. Hence, as in the static case, where the variable t does not appear, we have that this t-dependent resolvent matrix has a band block upper triangular structure

$$\omega^{[1]} = \begin{bmatrix} \omega^{[1]}_{0,0} & \omega^{[1]}_{0,1} & \omega^{[1]}_{0,2} & \dots & \omega^{[1]}_{0,N-1} & I_p & 0 & 0 & \dots \\ 0 & \omega^{[1]}_{1,1} & \omega^{[1]}_{1,2} & \dots & \omega^{[1]}_{1,N-1} & \omega^{[1]}_{1,N} & I_p & 0 & \dots \\ 0 & 0 & \omega^{[1]}_{2,2} & \dots & \omega^{[1]}_{2,N-1} & \omega^{[1]}_{2,N} & \omega^{[1]}_{2,N+1} & I_p & \ddots \\ & \ddots & \ddots & \ddots & & & & \ddots & \ddots \end{bmatrix}$$

with

$$\hat{H}_k(t) = \omega^{[1]}_{k,k}(t)H_k(t),$$

and the connection formulas described in Proposition 4.5 hold in this wider context.

Moreover, if $W(t)$ is a monic polynomial we can ensure that the Christoffel formula is also fulfilled for the non-Abelian 2D Toda and Theorem 4.11 remains valid also in this scenario. Formulas (4.8) and (4.10) hold directly and they do not need further explanation. However, (4.9) needs the following brief discussion. The Christoffel–Darboux kernel is defined exactly as we did in (1.13) there is no such a formula as the Christoffel Darboux formula given in Proposition 1.68 appear in this scenario. However, as was shown in [15], there are cases, such as the multigraded reductions, where one has a generalized Christoffel Darboux formula.

Chapter 5

Matrix Geronimus transformation for matrix bi-orthogonal polynomials on the real line

Geronimus transformations for orthogonal polynomials were first discussed in an exhaustive way in [76] in the framework of the Hahn's characterization of classical orthogonal polynomials, to find necessary and sufficient conditions in order to the sequences $(p_n(x))_{n\in\mathbb{N}}$ and $(\frac{p'_{n+1}(x)}{n+1})_{n\in\mathbb{N}}$ be orthogonal at the same time. There some determinantal formulas were found, (see also [149]). Now, our idea is to study the Geronimus transformation, but now, for a matrix of linear functionals $u=(u_{i,j})\in(O'_c)^{p\times p}$. The reader may observe certain similarities between the result of the original paper (see also Chapter 3) and this chapter.

5.1 Matrix Geronimus transformation

Definition 5.1. *Given a matrix polynomial $W(x)=\sum_{k=0}^{N}A_kx^k\in\mathbb{R}^{p\times p}[x]$ of degree N (not necessarily monic) and a matrix of linear functionals $u=(u_{i,j})\in(O'_c)^{p\times p}$, such that $\sigma(W)\cap\operatorname{supp}(u)=\varnothing$, a matrix of linear functionals $\check{u}$ is said to be a matrix Geronimus transformation of u, if*

$$\check{u}W(x)=u,$$

with the corresponding perturbed sesquilinear form satisfying

$$\left\langle P(x),Q(x)W(x)^{\dagger}\right\rangle_{\check{u}}=\langle P(x),Q(x)\rangle_{u}.$$

It is important to note that since $W(x)\in\mathbb{R}^{p\times p}[x]$, then $W(x)^{\dagger}=W(x)^{\top}$

Proposition 5.2. *The most general matrix Geronimus transformation is given by*

$$\check{u}:=u(W(x))^{-1}+v,\qquad v:=\sum_{a=1}^{q}\sum_{j=1}^{s_a}\sum_{m=0}^{\kappa_j^{(a)}-1}(-1)^m\delta^{(m)}(x-x_a)\frac{\xi_{j,m}^{[a]}}{m!}l_j^{(a)}(x),\tag{5.1}$$

where x_a, $a=1,\dots,q$, are the eigenvalues of $W(x)$ with their corresponding algebraic multiplicity $\kappa_j^{(a)}$ and v is expressed in terms of derivatives of Dirac linear functionals and adapted left root polynomials $l_j^{(a)}(x)$ of $W(x)$. Here $\xi_{j,m}^{[a]}\in\mathbb{C}^p$ are constant vectors.

Observe that v is associated with the eigenvalues and left root vectors of the perturbing polynomial $W(x)$. Notice that, when W has a singular leading coefficient, this spectral part could even disappear, for example if $W(x)$ is unimodular, i.e. with constant determinant, no depending on x. Observe that, in general, we have $Np\geq\sum_{a=1}^{q}\sum_{i=1}^{s_a}\kappa_j^{(a)}$ and we can not ensure the equality, up to for the nonsingular leading coefficient case. Let us assume that the perturbed moment matrix (associated with $\check{u}$) has a Gaussian factorization

$$\check{M}=\check{S}_1^{-1}\check{H}(\check{S}_2)^{-\dagger},$$

where $\check{S}_1,\check{S}_2$ are lower unitriangular block matrices and $\check{H}$ is a diagonal block matrix

$$\check{S}_i=\begin{bmatrix} I_p & 0_p & 0_p & \dots \\ (\check{S}_i)_{1,0} & I_p & 0_p & \cdots \\ (\check{S}_i)_{2,0} & (\check{S}_i)_{2,1} & I_p & \ddots \\ & & \ddots & \ddots \end{bmatrix}, \qquad i=1,2, \qquad \check{H}=\operatorname{diag}(\check{H}_0,\check{H}_1,\check{H}_2,\dots).$$

Hence, the perturbed matrix of linear functionals gives a family of matrix bi-orthogonal polynomials

$$\check{P}^{[i]}(x)=\check{S}_i\chi(x), \qquad i=1,2,$$

with respect to the perturbed sesquilinear form $\langle\cdot,\cdot\rangle_{\check{u}}$.

Proposition 5.3. *The moment matrices satisfy*

$$\check{M}W(\Lambda^\top)=M.$$

5.2 The resolvent and connection formulas

Definition 5.4. *The resolvent matrix is defined by*

$$\omega:=\check{S}_1(S_1)^{-1}. \tag{5.2}$$

Proposition 5.5. i) *The resolvent matrix can be also expressed as*

$$\omega=\check{H}\big(\check{S}_2\big)^{-\dagger}W(\Lambda^\top)\big(S_2\big)^\dagger H^{-1}. \tag{5.3}$$

ii) *The resolvent matrix is a lower unitriangular block matrix with at most the first N block subdiagonals different from zero, i.e.*

$$\omega = \begin{bmatrix} I_p & 0_p & \dots & 0_p & 0_p & \dots \\ \omega_{1,0} & I_p & \ddots & 0_p & 0_p & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ \omega_{N,0} & \omega_{N,1} & \dots & I_p & 0_p & \ddots \\ 0_p & \omega_{N+1,1} & \cdots & \omega_{N+1,N} & I_p & \ddots \\ \vdots & \ddots & \ddots & & \ddots & \ddots \end{bmatrix}.$$

iii) *The following connection formulas hold*

$$\check{P}^{[1]}(x) = \omega P^{[1]}(x), \tag{5.4}$$

$$\big(\check{H}^{-1}\omega H\big)^{\dagger}\check{P}^{[2]}(x) = P^{[2]}(x)W^{\dagger}(x). \tag{5.5}$$

iv) *For the last subdiagonal of the resolvent we have*

$$\omega_{N+k,k} = \check{H}_{N+k}A_N(H_k)^{-1}. \tag{5.6}$$

Proof. i) From Proposition 5.3 and the Gaussian factorization of M and $\check{M}$ we get

$$\big(S_1\big)^{-1}H\big(S_2\big)^{-\dagger} = \big(\check{S}_1\big)^{-1}\check{H}\big(\check{S}_2\big)^{-\dagger}W(\Lambda^{\top}),$$

and, consequently,

$$\check{S}_1\big(S_1\big)^{-1}H = \check{H}\big(\check{S}_2\big)^{-\dagger}W(\Lambda^{\top})\big(S_2\big)^{\dagger}.$$

From here the result follows.

ii) From the definition of the resolvent matrix, being a product of lower unitriangular matrices, it is also a lower unitriangular matrix. However, from (5.3) we deduce almost the contrary, we see that is a matrix with all its subdiagonals with zero coefficients but for the first N. Thus, it must have the described band structure.

iii) From the definition we have (5.4). Let us notice that (5.3) can be written as

$$\omega^{\dagger}\check{H}^{-\dagger} = H^{-\dagger}S_2W^{\dagger}(\Lambda)\big(\check{S}_2\big)^{-1},$$

so that

$$\omega^{\dagger}\check{H}^{-\dagger}\check{P}^{[2]}(x) = H^{-\dagger}S_2W^{\dagger}(\Lambda)\chi(x),$$

and (5.5) follows.

iv) It follows from (5.3).

∎

The connection formulas (5.4) and (5.5) can be written as

$$\check{P}_n^{[1]}(x) = P_n^{[1]}(x) + \sum_{k=\max\{0,n-N\}}^{n-1} \omega_{n,k} P_k^{[1]}(x), \tag{5.7}$$

$$W(x)\big(P_n^{[2]}(x)\big)^{\dagger}(H_n)^{-1} = \big(\check{P}_n^{[2]}(x)\big)^{\dagger}(\check{H}_n)^{-1} + \sum_{k=n+1}^{n+N} \big(\check{P}_k^{[2]}(x)\big)^{\dagger}(\check{H}_k)^{-1}\omega_{k,n}. \tag{5.8}$$

Lemma 5.6. *We have that*

$$\frac{W(x)-W(z)}{x-z} = [\chi(x)]_{[N]}^{\top}\mathcal{B}[\chi(z)]_{[N]}. \tag{5.9}$$

with $\mathcal{B}$ given in Definition 1.38.

Proof. It is a direct consequence of the fact that $\frac{x^n-z^n}{x-z} = \sum_{k=0}^{n-1} x^k z^{n-k-1}$ and

$$\begin{bmatrix} I_p & xI_p & \cdots & x^{N-1} \end{bmatrix} \begin{bmatrix} z^{N-1}I_p & z^{N-2}I_p & \cdots & \cdots & I_p \\ z^{N-2}I_p & z^{N-3}I_p & \cdots & I_p & \\ \vdots & \vdots & \iddots & & \\ zI_p & I_p & & & \\ I_p & & & & \end{bmatrix} \begin{bmatrix} A_N \\ \vdots \\ A_1 \end{bmatrix} = [\chi(x)]_{[N]}^{\top}\mathcal{B}[\chi(z)]_{[N]}.$$

∎

Proposition 5.7. *The Geronimus transformation of the function of second kind satisfies*

$$\omega C^{[1]}(z) = \check{C}^{[1]}(x)W(x) - \begin{bmatrix} \left(\check{H}\left(\check{S}_2\right)^{-\dagger}\right)_{[N]}\mathcal{B}(\chi(z))_{[N]} \\ 0_p \\ \vdots \end{bmatrix}, \tag{5.10}$$

and

$$\big(C^{[2]}(z)\big)^{\dagger}H^{-1} = \big(\check{C}^{[2]}(z)\big)^{\dagger}\check{H}^{-1}\omega. \tag{5.11}$$

Proof. From (1.9) we can write

$$\begin{aligned} \omega C^{[1]}(z) - \check{C}^{[1]}(z)W(z) &= \omega\left\langle P^{[1]}(x), \frac{I_p}{z-x}\right\rangle_u - \left\langle \check{P}^{[1]}(x), \frac{I_p}{z-x}\right\rangle_{\check{u}} W(z) \\ &= -\left\langle \check{P}^{[1]}(x), \frac{(W(x)-W(z))^{\dagger}}{x-z}\right\rangle_{\check{u}} = -\check{S}_1\left\langle \chi(x), \frac{(W(x)-W(z))^{\dagger}}{x-z}\right\rangle_{\check{u}}. \end{aligned}$$

Using the fact that $\frac{(W(x)-W(z))^{\dagger}}{x-z}$ is a matrix polynomial of degree $N-1$ and from (5.9), the result follows. Finally, in order to show (5.11), we note that from (5.8) and (1.9)

$$\big(C_n^{[2]}(z)\big)^{\dagger}(H_n)^{-1} = \big(\check{C}_n^{[2]}(z)\big)^{\dagger}(\check{H}_n)^{-1} + \sum_{k=n+1}^{n+N} \big(\check{C}_k^{[2]}(z)\big)^{\dagger}(\check{H}_k)^{-1}\omega_{k,n}.$$

∎

5.2.1 Connection formulas for perturbed Christoffel–Darboux kernels

Definition 5.8. *The resolvent wing is the matrix*

$$\Omega[n]=\begin{cases}\begin{bmatrix}\omega_{n,n-N} & \dots & \dots & \omega_{n,n-1}\\ 0_p & \ddots & & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ 0_p & \dots & 0_p & \omega_{n+N-1,n-1}\end{bmatrix}\in\mathbb{C}^{Np\times Np}, & n\geq N,\\ \begin{bmatrix}\omega_{n,0} & \dots & \dots & \omega_{n,n-1}\\ \vdots & & & \vdots\\ \omega_{N,0} & & & \omega_{N,n-1}\\ 0_p & \ddots & & \\ \vdots & \ddots & \ddots & \vdots\\ 0_p & \dots & 0_p & \omega_{n+N-1,n-1}\end{bmatrix}\in\mathbb{C}^{Np\times np}, & n<N.\end{cases}$$

Theorem 5.9. *For $m=\min(n,N)$, the perturbed and original Christoffel–Darboux kernels are related by the following connection formula*

$$\check{K}_{n-1}(x,y)=W(y)K_{n-1}(x,y)-\Big[\big(\check{P}^{[2]}_{n}(y)\big)^{\dagger}\check{H}_n^{-1},\dots,\big(\check{P}^{[2]}_{n+N-1}(y)\big)^{\dagger}\check{H}^{-1}_{n+N-1}\Big]\Omega[n]\begin{bmatrix}P^{[1]}_{n-m}(x)\\ \vdots\\ P^{[1]}_{n-1}(x)\end{bmatrix}. \tag{5.12}$$

For $n\geq N$, the connection formula for the mixed Christoffel–Darboux kernels reads

$$\check{K}^{(pc)}_{n-1}(x,y)W(x)=W(y)K^{(pc)}_{n-1}(x,y)-\Big[\big(\check{P}^{[2]}_{n}(y)\big)^{\dagger}\check{H}_n^{-1},\dots,\big(\check{P}^{[2]}_{n+N-1}(y)\big)^{\dagger}\check{H}^{-1}_{n+N-1}\Big]\Omega[n]\begin{bmatrix}C^{[1]}_{n-N}(x)\\ \vdots\\ C^{[1]}_{n-1}(x)\end{bmatrix}+\mathcal{V}(x,y), \tag{5.13}$$

where $\mathcal{V}(x,y)$ was introduced in Definition 1.42, see also Proposition 1.43.

Proof. For the first connection formulas (5.12) we consider

$$\mathcal{K}_{n-1}(x,y):=\Big[\big(\check{P}^{[2]}_{0}(y)\big)^{\dagger}(\check{H}_0)^{-1} \quad \cdots \quad \big(\check{P}^{[2]}_{n-1}(y)\big)^{\dagger}(\check{H}_{n-1})^{-1}\Big]\,\omega_{[n]}\begin{bmatrix}P^{[1]}_{0}(x)\\ \vdots\\ P^{[1]}_{n-1}(x)\end{bmatrix},$$

and we compute it in two different ways. From (5.7) we get that

$$\omega_{[n]} \begin{bmatrix} P_0^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x) \end{bmatrix} = \begin{bmatrix} \check{P}_0^{[1]}(x) \\ \vdots \\ \check{P}_{n-1}^{[1]}(x) \end{bmatrix},$$

and, therefore,

$$\mathcal{K}_{n-1}(x,y) = \check{K}_{n-1}(x,y).$$

Relation (5.8) leads to

$$\mathcal{K}_{n-1}(x,y) = W(y)K_{n-1}(x,y) - \left[\left(\check{P}_n^{[2]}(y)\right)^{\dagger}(\check{H}_n)^{-1},\dots,\left(\check{P}_{n+N-1}^{[2]}(y)\right)^{\dagger}(\check{H}_{n+N-1})^{-1} \right] \Omega[n] \begin{bmatrix} P_{n-m}^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x) \end{bmatrix},$$

and the formula follows. To prove (5.13) we now consider

$$\mathcal{K}_{n-1}^{(pc)}(x,y) := \left[\left(\check{P}_0^{[2]}(y)\right)^{\dagger}(\check{H}_0)^{-1},\dots,\left(\check{P}_{n-1}^{[2]}(y)\right)^{\dagger}(\check{H}_{n-1})^{-1} \right] \omega_{[n]} \begin{bmatrix} C_0^{[1]}(x) \\ \vdots \\ C_{n-1}^{[1]}(x) \end{bmatrix},$$

which, as above, can be computed in two different forms. In one hand, using (5.10) we get

$$\begin{aligned} \mathcal{K}_{n-1}^{(pc)}(x,y) &= \left[\left(\check{P}_0^{[2]}(y)\right)^{\dagger}(\check{H}_0)^{-1},\dots,\left(\check{P}_{n-1}^{[2]}(y)\right)^{\dagger}(\check{H}_{n-1})^{-1} \right] \left(\begin{bmatrix} \check{C}_0^{[1]}(x)W(x) \\ \vdots \\ \check{C}_{n-1}^{[1]}(x)W(x) \end{bmatrix} - \left(\check{H}\left(\check{S}_2\right)^{-\dagger}\right)_{[n,N]} \mathcal{B}(\chi(z))_{[N]} \right) \\ &= \check{K}_{n-1}^{(pc)}(x,y)W(x) - \left((\chi(y))_{[n]}\right)^{\dagger} \left(\left(\check{S}_2\right)^{\dagger}\check{H}^{-1}\right)_{[n]} \left(\check{H}\left(\check{S}_2\right)^{-\dagger}\right)_{[n,N]} \mathcal{B}(\chi(x))_{[N]}, \end{aligned}$$

where $\left(\check{H}\left(\check{S}_2\right)^{-\dagger}\right)_{[n,N]}$ is the truncation to the n first block rows and first N block columns of $\check{H}\left(\check{S}_2\right)^{-\dagger}$. This simplifies for $n \geq N$ to

$$\mathcal{K}_{n-1}^{(pc)}(x,y) = \check{K}_{n-1}^{(pc)}(x,y)W(x) - \left((\chi(y))_{[N]}\right)^{\dagger} \mathcal{B}(\chi(x))_{[N]}.$$

On the other hand, if we use (5.8),

$$\mathcal{K}_{n-1}^{(pc)}(x,y) = W(y)K_{n-1}^{(pc)}(x,y) - \left[\left(\check{P}_n^{[2]}(y)\right)^{\dagger}(\check{H}_n)^{-1},\dots,\left(\check{P}_{n+N-1}^{[2]}(y)\right)^{\dagger}(\check{H}_{n+N-1})^{-1} \right] \Omega[n] \begin{bmatrix} C_{n-N}^{[1]}(x) \\ \vdots \\ C_{n-1}^{[1]}(x) \end{bmatrix},$$

and, consequently, for $n \geq N$ we obtain

$$\check{K}^{(pc)}_{n-1}(x,y)W(x) = W(y)K^{(pc)}_{n-1}(x,y) - \Big[\big(\check{P}^{[2]}_{n}(y)\big)^{\dagger}(\check{H}_n)^{-1},\dots,\big(\check{P}^{[2]}_{n+N-1}(y)\big)^{\dagger}(\check{H}_{n+N-1})^{-1}\Big]\Omega[n]\begin{bmatrix} C^{[1]}_{n-N}(x) \\ \vdots \\ C^{[1]}_{n-1}(x)\end{bmatrix}$$
$$+\big((\chi(y))_{[N]}\big)^{\dagger}\mathcal{B}(\chi(z))_{[N]}.$$

■

5.2.2 Spectral properties of the first family of perturbed second kind functions

From now on we will assume that the perturbing polynomial is monic, i.e. $W(x) = I_p x^N + \sum_{k=0}^{N-1} A_k x^k \in \mathbb{R}^{p\times p}[x]$. Thus, we can apply the spectral theory of this polynomial (see Chapter 1 as well as [77]).

Definition 5.10. *Let us introduce two upper block triangular matrices. First, in terms of the couplings of masses in the Geronimus perturbation (see (5.1))*

$$\mathcal{X}^{(a)}_i := \begin{bmatrix} \xi^{[a]}_{i,\kappa^{(a)}_i-1} & \xi^{[a]}_{i,\kappa^{(a)}_i-2} & \xi^{[a]}_{i,\kappa^{(a)}_i-3} & \dots & \xi^{[a]}_{i,0} \\ 0_{p\times 1} & \xi^{[a]}_{i,\kappa^{(a)}_i-1} & \xi^{[a]}_{i,\kappa^{(a)}_i-2} & \dots & \xi^{[a]}_{i,1} \\ 0_{p\times 1} & 0_{p\times 1} & \xi^{[a]}_{i,\kappa^{(a)}_i-1} & \dots & \xi^{[a]}_{i,2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{p\times 1} & 0_{p\times 1} & & & \xi^{[a]}_{i,\kappa^{(a)}_i-1}\end{bmatrix} \in \mathbb{C}^{p\kappa^{(a)}_i\times\kappa^{(a)}_i}. \tag{5.14}$$

In terms of the left Jordan chains we also define

$$\mathcal{L}^{(a)}_i := \begin{bmatrix} l^{(a)}_{i,0} & l^{(a)}_{i,1} & l^{(a)}_{i,2} & \dots & l^{(a)}_{i,\kappa^{(a)}_i-1} \\ 0_{1\times p} & l^{(a)}_{i,0} & l^{(a)}_{i,1} & \dots & l^{(a)}_{i,\kappa^{(a)}_i-2} \\ 0_{1\times p} & 0_{1\times p} & l^{(a)}_{i,0} & \dots & l^{(a)}_{i,\kappa^{(a)}_i-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{1\times p} & 0_{1\times p} & & & l^{(a)}_{i,0}\end{bmatrix} \in \mathbb{C}^{\kappa^{(a)}_i\times p\kappa^{(a)}_i}. \tag{5.15}$$

For $z \neq x_a$, we introduce the $p\times p$ matrices

$$\check{C}^{(a)}_{n;i}(z) := \mathbf{J}^{(i)}_{\check{P}^{[1]}_n}(x_a)\mathcal{X}^{(a)}_i\mathcal{L}^{(a)}_i \begin{bmatrix} \frac{I_p}{(z-x_a)^{\kappa^{(a)}_i}} \\ \vdots \\ \frac{I_p}{z-x_a}\end{bmatrix}, \tag{5.16}$$

where $i = 1, \dots, s_a$, and the matrix $\mathtt{J}_f^{(i)}$ was given in Definition 1.35 .

Observe that $\mathcal{X}_i^{(a)} \mathcal{L}_i^{(a)} \in \mathbb{C}^{p\kappa_i^{(a)} \times p\kappa_i^{(a)}}$ is a block upper triangular matrix, with blocks in $\mathbb{C}^{p\times p}$.

Proposition 5.11. *For $z \notin \operatorname{supp}(u) \cup \sigma(W)$, the following expression*

$$\check{C}_n^{[1]}(z) = \left\langle \check{P}_n^{[1]}(x), \frac{I_p}{z-x} \right\rangle_{uW^{-1}} + \sum_{a=1}^{q} \sum_{i=1}^{s_a} \check{C}_{n;i}^{(a)}(z)$$

holds.

Proof. From the definition of function of second kind and Geronimus transformation

$$\begin{aligned} \check{C}_n^{[1]}(z) &= \left\langle \check{P}_n^{[1]}(x), \frac{I_p}{z-x} \right\rangle_{\check{u}} \\ &= \left\langle \check{P}_n^{[1]}(x), \frac{I_p}{z-x} \right\rangle_{uW^{-1}} + \sum_{a=1}^{q} \sum_{i=1}^{s_a} \sum_{m=0}^{\kappa_i^{(a)}-1} \left(\check{P}_n^{[1]}(x) \frac{\xi_{i,m}^{[a]}}{m!} \frac{l_i^{(a)}(x)}{z-x} \right)^{(m)}_{x=x_a} . \end{aligned}$$

Now, taking into account that

$$\left(\check{P}_n^{[1]}(x) \frac{\xi_{i,m}^{[a]}}{m!} \frac{l_i^{(a)}(x)}{z-x} \right)^{(m)}_{x=x_a} = \sum_{k=0}^{m} \left(\check{P}_n^{[1]}(x) \frac{\xi_{i,m}^{[a]} l_i^{(a)}(x)}{(m-k)!} \right)^{(m-k)}_{x=x_a} \frac{1}{(z-x_a)^{k+1}},$$

we deduce the result. ■

Lemma 5.12. *Let $r_j^{(a)}(x)$ be the adapted right root polynomials of the monic matrix polynomial $W(x)$ given in (1.3). Then*

$$\mathcal{L}_i^{(a)} \begin{bmatrix} \dfrac{I_p}{(x-x_a)^{\kappa_i^{(a)}}} \\ \vdots \\ \dfrac{I_p}{x-x_a} \end{bmatrix} W(x) r_j^{(a)}(x) = \begin{bmatrix} \dfrac{1}{(x-x_a)^{\kappa_i^{(a)}}} \\ \vdots \\ \dfrac{1}{x-x_a} \end{bmatrix} l_i^{(a)}(x) W(x) r_j^{(a)}(x) + (x-x_a)^{\kappa_j^{(a)}} T(x), \quad T(x) \in \mathbb{C}^{\kappa_j^{(a)}}[x].$$

Proof. Notice that we can write

$$\mathcal{L}_i^{(a)} \begin{bmatrix} \dfrac{I_p}{(x-x_a)^{\kappa_i^{(a)}}} \\ \vdots \\ \dfrac{I_p}{x-x_a} \end{bmatrix} W(x) r_j^{(a)}(x) = \begin{bmatrix} l_{i,0}^{(a)} & l_{i,1}^{(a)} & l_{i,2}^{(a)} & \cdots & l_{i,\kappa_i^{(a)}-1}^{(a)} \\ 0_{1\times p} & l_{i,0}^{(a)} & l_{i,1}^{(a)} & \cdots & l_{i,\kappa_i^{(a)}-2}^{(a)} \\ 0_{1\times p} & 0_{1\times p} & l_{i,0}^{(a)} & & l_{i,\kappa_i^{(a)}-3}^{(a)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{1\times p} & 0_{1\times p} & & & l_{i,0}^{(a)} \end{bmatrix} \begin{bmatrix} \dfrac{I_p}{(x-x_a)^{\kappa_i^{(a)}}} \\ \vdots \\ \dfrac{I_p}{x-x_a} \end{bmatrix} W(x) r_j^{(a)}(x)$$

$$
= \begin{bmatrix} \dfrac{l_i^{(a)}(x)}{(x-x_a)^{\kappa_i^{(a)}}} \\ \dfrac{l_i^{(a)}(x)}{(x-x_a)^{\kappa_i^{(a)}-1}} - l_{i,\kappa_i^{(a)}-1}^{(a)} \\ \vdots \\ \dfrac{l_i^{(a)}(x)}{x-x_a} - l_{i,1}^{(a)} - \cdots - l_{i,\kappa_i^{(a)}-1}^{(a)}(x-x_a)^{\kappa_i^{(a)}-2} \end{bmatrix} W(x) r_j^{(a)}(x).
$$

Now, (1.5) yields the result. ∎

Lemma 5.13. *The function $\check{C}_{n;i}^{(a)}(x)W(x)r_j^{(b)}(x) \in \mathbb{C}^p[x]$ satisfies*

$$
\check{C}_{n;i}^{(a)}(x)W(x)r_j^{(b)}(x) = \begin{cases} \mathbf{J}_{\check{P}_n^{[1]}}^{(i)}(x_a)\mathcal{X}_i^{(a)} \begin{bmatrix} (x-x_a)^{\kappa_{\max(i,j)}^{(a)}-\kappa_i^{(a)}} \\ \vdots \\ (x-x_a)^{\kappa_{\max(i,j)}^{(a)}-1} \end{bmatrix} w_{i,j}^{(a)}(x) + (x-x_a)^{\kappa_j^{(a)}} T^{(a,a)}(x), & \text{if } a = b, \\ (x-x_b)^{\kappa_j^{(b)}} T^{(a,b)}(x), & \text{if } a \neq b, \end{cases} \tag{5.17}
$$

where the $\mathbb{C}^p$-valued function $T^{(a,b)}(x)$ is analytic at $x = x_b$, and, in particular, $T^{(a,a)}(x) \in \mathbb{C}^p[x]$.

Proof. First, for the function $\check{C}_{n;i}^{(a)}(x)W(x)r_j^{(b)}(x) \in \mathbb{C}^p[x]$, with $a \neq b$, and recalling (1.5), we have

$$
\begin{aligned}
\check{C}_{n;i}^{(a)}(x)W(x)r_j^{(b)}(x) &= \mathbf{J}_{\check{P}_n^{[1]}}^{(i)}(x_a)\mathcal{X}_i^{(a)}\mathcal{L}_i^{(a)} \begin{bmatrix} \dfrac{I_p}{(x-x_a)^{\kappa_i^{(a)}}} \\ \vdots \\ \dfrac{I_p}{x-x_a} \end{bmatrix} W(x)r_j^{(b)}(x) \\
&= (x-x_b)^{\kappa_j^{(b)}} T^{(a,b)}(x),
\end{aligned}
$$

where the $\mathbb{C}^p$-valued function $T^{(a,b)}(x)$ is analytic at $x = x_b$. Second, from (5.16) and Lemma 5.12 we deduce that

$$
\check{C}_{n;i}^{(a)}(x)W(x)r_j^{(a)}(x) = \mathbf{J}_{\check{P}_n^{[1]}}^{(i)}(x_a)\mathcal{X}_i^{(a)}\mathcal{L}_i^{(a)} \begin{bmatrix} \dfrac{I_p}{(x-x_a)^{\kappa_i^{(a)}}} \\ \vdots \\ \dfrac{I_p}{x-x_a} \end{bmatrix} W(x)r_j^{(a)}(x)
$$

$$
\begin{aligned}
&= \mathtt{J}^{(i)}_{\check{P}^{[1]}_n}(x_a)\mathcal{X}^{(a)}_i \begin{bmatrix} \dfrac{1}{(x-x_a)^{\kappa^{(a)}_i}} \\ \vdots \\ \dfrac{1}{x-x_a} \end{bmatrix} l^{(a)}_i(x)W(x)r^{(a)}_j(x) \\
&\qquad + (x-x_a)^{\kappa^{(a)}_j}\mathtt{J}^{(i)}_{\check{P}^{[1]}_n}(x_a)\mathcal{X}^{(a)}_i T^{(a,a)}(x),
\end{aligned}
$$

for some $T^{(a,a)}(x) \in \mathbb{C}^p[x]$. Therefore, from Proposition 1.34 we get

$$
\begin{aligned}
\check{C}^{(a)}_{n;i}(x)W(x)r^{(a)}_j(x) &= \mathtt{J}^{(i)}_{\check{P}^{[1]}_n}(x_a)\mathcal{X}^{(a)}_i \begin{bmatrix} (x-x_a)^{\kappa^{(a)}_{\max(i,j)}-\kappa^{(a)}_i} \\ \vdots \\ (x-x_a)^{\kappa^{(a)}_{\max(i,j)}-1} \end{bmatrix} \\
&\quad \times \left(w^{(a)}_{i,j;0} + w^{(a)}_{i,j;1}(x-x_a) + \cdots + w^{(a)}_{i,j;\kappa^{(a)}_{\min(i,j)}+N-2}(x-x_a)^{\kappa^{(a)}_{\min(i,j)}+N-2} \right). \\
&\qquad + (x-x_a)^{\kappa^{(a)}_j}\mathtt{J}^{(i)}_{\check{P}^{[1]}_n}(x_a)\mathcal{X}^{(a)}_i T^{(a,a)}(x),
\end{aligned}
$$

and the result follows. ∎

Lemma 5.14. *For $m = 0, \dots, \kappa^{(a)}_j - 1$, the following relations hold*

$$
\left(\check{C}^{[1]}_n(z)W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a} = \sum_{i=1}^{s_a} \left(\check{C}^{(a)}_{n;i}(z)W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a}. \tag{5.18}
$$

Proof. For $z \notin \operatorname{supp}(u) \cup \sigma(W)$, Proposition 5.11 leads to

$$
\left(\check{C}^{[1]}_n(z)W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a} = \left(\left\langle \check{P}^{[1]}_n(x), \frac{I_p}{z-x}\right\rangle_{uW^{-1}} W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a} + \sum_{b=1}^{q}\sum_{i=1}^{s_b}\left(\check{C}^{(b)}_{n;i}(z)W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a}.
$$

But, as $\sigma(W) \cap \operatorname{supp}(u) = \varnothing$, the derivatives of the Cauchy kernel $1/(z-x)$ are analytic functions at $z = x_a$. Consequently,

$$
\begin{aligned}
\left(\left\langle \check{P}^{[1]}_n(x), \frac{I_p}{z-x}\right\rangle_{uW^{-1}} W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a} &= \left\langle \check{P}^{[1]}_n(x), \left(\frac{[W(z)r^{(a)}_j(z)]^{\dagger}}{z-x}\right)^{(m)}_{z=x_a}\right\rangle_{uW^{-1}} \\
&= \left\langle \check{P}^{[1]}_n(x), \left[\sum_{k=0}^{m}\binom{m}{k}\left(W(z)r^{(a)}_j(z)\right)^{(k)}_{z=x_a}\frac{(-1)^{m-k}(m-k)!}{(x_a-x)^{m-k+1}}\right]^{\dagger}\right\rangle_{uW^{-1}} \\
&= 0_{p\times 1},
\end{aligned}
$$

for $m = 0, \dots, \kappa^{(a)}_j - 1$, where in the last equation we have used (1.5). Equation (5.17) shows that $\check{C}^{(b)}_{n;i}(z)W(z)r^{(a)}_j(x)$ for $b \neq a$ has a zero at $z = x_a$ of order $\kappa^{(a)}_j$ and, as a consequence,

$$
\left(\check{C}^{(b)}_{n;i}(z)W(z)r^{(a)}_j(z)\right)^{(m)}_{z=x_a} = 0, \quad b \neq a, \quad \text{for} \quad m = 0, \dots, \kappa^{(a)}_j - 1.
$$

■

Definition 5.15. *Let us introduce the matrix* $\mathcal{W}_{j,i}^{(a)} \in \mathbb{C}^{\kappa_j^{(a)} \times \kappa_i^{(a)}}$

$$\mathcal{W}_{j,i}^{(a)} := \begin{cases} \left[\begin{array}{c|cccc} & w_{i,j;0}^{(a)} & w_{i,j;1}^{(a)} & \cdots & w_{i,j;\kappa_j^{(a)}-1}^{(a)} \\ & 0 & w_{i,j;0}^{(a)} & \cdots & w_{i,j;\kappa_j^{(a)}-2}^{(a)} \\ 0_{\kappa_j^{(a)}\times(\kappa_i^{(a)}-\kappa_j^{(a)})} & \vdots & & \ddots & \vdots \\ & 0 & 0 & & w_{i,j;0}^{(a)} \end{array}\right], & i \geq j, \\ \begin{bmatrix} w_{i,j;\kappa_j^{(a)}-\kappa_i^{(a)}}^{(a)} & w_{i,j;\kappa_j^{(a)}-\kappa_i^{(a)}+1}^{(a)} & \cdots & w_{i,j,\kappa_j^{(a)}-1}^{(a)} \\ \vdots & & & \vdots \\ w_{i,j;0}^{(a)} & w_{i,j;1}^{(a)} & \cdots & w_{i,j;\kappa_i^{(a)}-1}^{(a)} \\ 0 & w_{i,j;0}^{(a)} & & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & & w_{i,j;0}^{(a)} \end{bmatrix}, & i \leq j, \end{cases}$$

and the matrix $\mathcal{W}_j^{(a)} \in \mathbb{C}^{\kappa_j^{(a)} \times \alpha_a}$ *given by*

$$\mathcal{W}_j^{(a)} := \left[\mathcal{W}_{j,1}^{(a)}, \dots, \mathcal{W}_{j,s_a}^{(a)}\right].$$

We also consider the matrices $\mathcal{T}_i^{(a)} \in \mathbb{C}^{p\kappa_i^{(a)} \times \alpha_a}$, $\mathcal{T}^{(a)} \in \mathbb{C}^{p\alpha_a \times \alpha_a}$, *and* $\mathcal{T} \in \mathbb{C}^{Np^2 \times Np}$

$$\mathcal{T}_i^{(a)} := \mathcal{X}_i^{(a)} \mathcal{W}_i^{(a)}, \qquad \mathcal{T}^{(a)} := \begin{bmatrix} \mathcal{T}_1^{(a)} \\ \vdots \\ \mathcal{T}_{s_a}^{(a)} \end{bmatrix}, \qquad \mathcal{T} := \operatorname{diag}(\mathcal{T}^{(1)}, \dots, \mathcal{T}^{(q)}). \tag{5.19}$$

Proposition 5.16. *The equations*

$$\mathcal{J}^{(j)}_{\check{C}_n^{[1]}W}(x_a) = \sum_{i=1}^{s_a} \mathcal{J}^{(j)}_{\check{C}_{n;i}^{(a)}W}(x_a), \qquad \mathcal{J}_{\check{C}_n^{[1]}W}(x_a) = \sum_{i=1}^{s_a} \mathcal{J}_{\check{C}_{n;i}^{(a)}W}(x_a), \tag{5.20}$$

$$\mathcal{J}^{(j)}_{\check{C}_{n;i}^{(a)}W}(x_a) = \mathbf{J}^{(i)}_{\check{P}_n^{[1]}}(x_a)\mathcal{X}_i^{(a)}\mathcal{W}_{i,j}^{(a)}, \qquad \mathcal{J}_{\check{C}_{n;i}^{(a)}W}(x_a) = \mathbf{J}^{(i)}_{\check{P}_n^{[1]}}(x_a)\mathcal{X}_i^{(a)}\mathcal{W}_i^{(a)}, \tag{5.21}$$

$$\mathcal{J}_{\check{C}_n^{[1]}W}(x_a) = \mathbf{J}_{\check{P}_n^{[1]}}(x_a)\mathcal{T}^{(a)}, \qquad \mathcal{J}_{\check{C}_n^{[1]}W} = \mathbf{J}_{\check{P}_n^{[1]}}\mathcal{T}, \tag{5.22}$$

hold.

Proof. Eq. (5.20) is a direct consequence of (5.18). According to (5.17) for $m = 0, \dots, \kappa_j^{(a)} - 1$, we have

$$\big(\check{C}_{n;i}^{(a)}(x)W(x)r_j^{(a)}(x)\big)_{x=x_a}^{(m)} = \mathtt{J}^{(i)}_{\check{P}_n^{[1]}}(x_a)\mathcal{X}_i^{(a)}\begin{bmatrix}\Big((x-x_a)^{\kappa^{(a)}_{\max(i,j)}-\kappa_i^{(a)}}w_{i,j}^{(a)}(x)\Big)_{x=x_a}^{(m)}\\ \vdots \\ \Big((x-x_a)^{\kappa^{(a)}_{\max(i,j)}-1}w_{i,j}^{(a)}(x)\Big)_{x=x_a}^{(m)}\end{bmatrix},$$

and collecting all these equations in a matrix form we get (5.21). Finally, for (5.22) notice that from (5.20) and (5.21) we deduce

$$\mathcal{J}^{(j)}_{\check{C}_n^{[1]}W}(x_a) = \sum_{i=1}^{s_a}\mathtt{J}^{(i)}_{\check{P}_n^{[1]}}(x_a)\mathcal{X}_i^{(a)}\mathcal{W}_{i,j}^{(a)}, \qquad \mathcal{J}_{\check{C}_n^{[1]}W}(x_a) = \sum_{i=1}^{s_a}\mathtt{J}^{(i)}_{\check{P}_n^{[1]}}(x_a)\mathcal{X}_i^{(a)}\mathcal{W}_i^{(a)}.$$

Now, using (5.19) we can write the second equation as follows

$$\begin{aligned}\mathcal{J}_{\check{C}_n^{[1]}W}(x_a) &= \sum_{i=1}^{s_a}\mathtt{J}^{(i)}_{\check{P}_n^{[1]}}(x_a)\mathcal{T}_i^{(a)}\\ &= \mathtt{J}_{\check{P}_n^{[1]}}(x_a)\mathcal{T}^{(a)}.\end{aligned}$$

A similar argument yields the second relation. ■

5.2.3 Spectral Christoffel–Geronimus formulas

Discussion for $n \geq N$.

Remark 5.17. *We stress that later on (see Corollary 5.38), using a nonspectral approach, we will see that*

$$\det\begin{bmatrix}\mathcal{J}_{C_{n-N}^{[1]}} - \mathtt{J}_{P_{n-N}^{[1]}}\mathcal{T}\\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}} - \mathtt{J}_{P_{n-1}^{[1]}}\mathcal{T}\end{bmatrix} \neq 0, \qquad n \geq N.$$

Proposition 5.18. *For $n \geq N$, the matrix coefficients of the connection matrix satisfy*

$$\big[\omega_{n,n-N}, \dots, \omega_{n,n-1}\big] = -\big(\mathcal{J}_{C_n^{[1]}} - \mathtt{J}_{P_n^{[1]}}\mathcal{T}\big)\begin{bmatrix}\mathcal{J}_{C_{n-N}^{[1]}} - \mathtt{J}_{P_{n-N}^{[1]}}\mathcal{T}\\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}} - \mathtt{J}_{P_{n-1}^{[1]}}\mathcal{T}\end{bmatrix}^{-1}.$$

Proof. From the connection formula (5.10), for $n \geq N$

$$\check{C}_n^{[1]}(x)W(x) = \sum_{k=n-N}^{n-1}\omega_{n,k}C_k^{[1]}(x) + C_n^{[1]}(x),$$

and we conclude that

$$\mathcal{J}_{\check{C}_n^{[1]}W} = \begin{bmatrix}\omega_{n,n-N},\dots,\omega_{n,n-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{C_{n-N}^{[1]}}\\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}}\end{bmatrix} + \mathcal{J}_{C_n^{[1]}}.$$

Similarly, using (5.7), we get

$$\mathtt{J}_{\check{P}_n^{[1]}} = \begin{bmatrix}\omega_{n,n-N},\dots,\omega_{n,n-1}\end{bmatrix}\begin{bmatrix}\mathtt{J}_{P_{n-N}^{[1]}}\\ \vdots \\ \mathtt{J}_{P_{n-1}^{[1]}}\end{bmatrix} + \mathtt{J}_{P_n^{[1]}}. \tag{5.23}$$

Now, from (5.22) we deduce

$$\begin{bmatrix}\omega_{n,n-N},\dots,\omega_{n,n-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{C_{n-N}^{[1]}}\\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}}\end{bmatrix} + \mathcal{J}_{C_n^{[1]}} = \begin{bmatrix}\omega_{n,n-N},\dots,\omega_{n,n-1}\end{bmatrix}\begin{bmatrix}\mathtt{J}_{P_{n-N}^{[1]}}\mathcal{T}\\ \vdots \\ \mathtt{J}_{P_{n-1}^{[1]}}\mathcal{T}\end{bmatrix} + \mathtt{J}_{P_n^{[1]}}\mathcal{T},$$

or, equivalently

$$\begin{bmatrix}\omega_{n,n-N},\dots,\omega_{n,n-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{C_{n-N}^{[1]}} - \mathtt{J}_{P_{n-N}^{[1]}}\mathcal{T}\\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}} - \mathtt{J}_{P_{n-1}^{[1]}}\mathcal{T}\end{bmatrix} = -\left(\mathcal{J}_{C_n^{[1]}} - \mathtt{J}_{P_n^{[1]}}\mathcal{T}\right).$$

Hence, using Remark 5.17, we get the result. ■

Remark 5.19. *In the next theorem we will need a spectral root jet vector but for the Christoffel–Darboux kernel and for its mixed version. In these kernels we have two variables, x and y. The jets are taken with respect to the first variable x, and we treat the y-variable as a parameter. We denote these matrix spectral jets of the Christoffel–Darboux kernel by $\mathtt{J}_{K_n}(y)$ and the root spectral jet vector of the mixed Christoffel–Darboux kernel by $\mathcal{J}_{K_n^{(pc)}}(y)$.*

Theorem 5.20 (Spectral Christoffel–Geronimus formulas)**.** *Assuming $n \geq N$, for monic perturbations of the matrix of linear functionals we have the following last quasideterminantal expressions for the perturbed bi-orthogonal matrix polynomials and its matrix norms*

$$\check{P}_n^{[1]}(x) = \Theta_* \begin{bmatrix}\mathcal{J}_{C_{n-N}^{[1]}} - \mathtt{J}_{P_{n-N}^{[1]}}\mathcal{T} & P_{n-N}^{[1]}(x)\\ \vdots & \vdots \\ \mathcal{J}_{C_n^{[1]}} - \mathtt{J}_{P_n^{[1]}}\mathcal{T} & P_n^{[1]}(x)\end{bmatrix},$$

$$\check{H}_n=\Theta_*\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T} & H_{n-N}\\ \mathcal{J}_{C^{[1]}_{n-N+1}}-\mathbf{J}_{P^{[1]}_{n-N+1}}\mathcal{T} & 0_p\\ \vdots & \vdots\\ \mathcal{J}_{C^{[1]}_{n}}-\mathbf{J}_{P^{[1]}_{n}}\mathcal{T} & 0_p\end{bmatrix},$$

$$\big(\check{P}^{[2]}_n(y)\big)^\dagger=-\Theta_*\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T} & H_{n-N}\\ \mathcal{J}_{C^{[1]}_{n-N+1}}-\mathbf{J}_{P^{[1]}_{n-N+1}}\mathcal{T} & 0_p\\ \vdots & \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}}-\mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T} & 0_p\\ W(y)\big(\mathcal{J}_{K^{(pc)}_{n-1}}(y)-\mathbf{J}_{K_{n-1}}(y)\mathcal{T}\big)+\mathcal{J}_{\mathcal{V}}(y) & 0_p\end{bmatrix}.$$

Proof. First, we consider the expressions for the family $\check{P}^{[1]}_n(x)$ and the quasitau matrices $\check{H}_n$. Using (5.7) we have

$$\check{P}^{[1]}_n(x)=P^{[1]}_n(x)+\big[\omega_{n,n-N},\dots,\omega_{n,n-1}\big]\begin{bmatrix}P^{[1]}_{n-N}(x)\\ \vdots\\ P^{[1]}_{n-1}(x)\end{bmatrix},$$

and from Proposition 5.18 we obtain

$$\check{P}^{[1]}_n(x)=P^{[1]}_n(x)-\big(\mathcal{J}_{C^{[1]}_{n}}-\mathbf{J}_{P^{[1]}_{n}}\mathcal{T}\big)\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}}-\mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix}^{-1}\begin{bmatrix}P^{[1]}_{n-N}(x)\\ \vdots\\ P^{[1]}_{n-1}(x)\end{bmatrix},$$

and the result follows. To get the transformation for the H's we proceed as follows. From (5.6) we deduce

$$\check{H}_n=\omega_{n,n-N}H_{n-N}.$$

But, according to Proposition 5.18, we have

$$\omega_{n,n-N}=-\big(\mathcal{J}_{C^{[1]}_{n}}-\mathbf{J}_{P^{[1]}_{n}}\mathcal{T}\big)\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}}-\mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix}^{-1}\begin{bmatrix}I_p\\ 0_p\\ \vdots\\ 0_p\end{bmatrix}.$$

Hence,

$$\check{H}_n=-\big(\mathcal{J}_{C^{[1]}_{n}}-\mathbf{J}_{P^{[1]}_{n}}\mathcal{T}\big)\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}}-\mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix}^{-1}\begin{bmatrix}H_{n-N}\\ 0_p\\ \vdots\\ 0_p\end{bmatrix},$$

and the results follows. We now prove the result for the second family. In one hand, according to Definition 1.13 we rewrite (5.13) as

$$\sum_{k=0}^{n-1}\big(\check P_k^{[2]}(y)\big)^{\dagger}\check H_k^{-1}\check C_k^{[1]}(x)W(x)=W(y)K_{n-1}^{(pc)}(x,y)\\ -\Big[\big(\check P_n^{[2]}(y)\big)^{\dagger}\check H_n^{-1},\dots,\big(\check P_{n+N-1}^{[2]}(y)\big)^{\dagger}\check H_{n+N-1}^{-1}\Big]\Omega[n]\begin{bmatrix}C_{n-N}^{[1]}(x)\\ \vdots\\ C_{n-1}^{[1]}(x)\end{bmatrix}+\mathcal V(x,y).$$

Therefore, the corresponding spectral jets satisfy

$$\sum_{k=0}^{n-1}\big(\check P_k^{[2]}(y)\big)^{\dagger}\check H_k^{-1}\mathcal J_{\check C_k^{[1]}W}=W(y)\mathcal J_{K_{n-1}^{(pc)}}(y)\\ -\Big[\big(\check P_n^{[2]}(y)\big)^{\dagger}\check H_n^{-1},\dots,\big(\check P_{n+N-1}^{[2]}(y)\big)^{\dagger}\check H_{n+N-1}^{-1}\Big]\Omega[n]\begin{bmatrix}\mathcal J_{C_{n-N}^{[1]}}\\ \vdots\\ \mathcal J_{C_{n-1}^{[1]}}\end{bmatrix}+\mathcal J_{\mathcal V}(y),$$

and, recalling (5.22), we conclude that

$$\sum_{k=0}^{n-1}\big(\check P_k^{[2]}(y)\big)^{\dagger}\check H_k^{-1}\mathbf J_{\check P_k^{[1]}}\mathcal T=W(y)\mathcal J_{K_{n-1}^{(pc)}}(y)\\ -\Big[\big(\check P_n^{[2]}(y)\big)^{\dagger}\check H_n^{-1},\dots,\big(\check P_{n+N-1}^{[2]}(y)\big)^{\dagger}\check H_{n+N-1}^{-1}\Big]\Omega[n]\begin{bmatrix}\mathcal J_{C_{n-N}^{[1]}}\\ \vdots\\ \mathcal J_{C_{n-1}^{[1]}}\end{bmatrix}+\mathcal J_{\mathcal V}(y).\tag{5.24}$$

On the other hand, from (5.12) we realize that

$$\sum_{k=0}^{n-1}\big(\check P_k^{[2]}(y)\big)^{\dagger}\check H_k^{-1}\mathbf J_{\check P_k^{[1]}}\mathcal T=W(y)\mathbf J_{K_{n-1}}(y)\mathcal T-\Big[\big(\check P_n^{[2]}(y)\big)^{\dagger}\check H_n^{-1},\dots,\big(\check P_{n+N-1}^{[2]}(y)\big)^{\dagger}\check H_{n+N-1}^{-1}\Big]\Omega[n]\begin{bmatrix}\mathbf J_{P_{n-N}^{[1]}}\mathcal T\\ \vdots\\ \mathbf J_{P_{n-1}^{[1]}}\mathcal T\end{bmatrix},$$

which can be subtracted to (5.24) to get

$$W(y)\big(\mathcal J_{K_{n-1}^{(pc)}}(y)-\mathbf J_{K_{n-1}}(y)\mathcal T\big)+\mathcal J_{\mathcal V}(y)=\Big[\big(\check P_n^{[2]}(y)\big)^{\dagger}\check H_n^{-1},\dots,\big(\check P_{n+N-1}^{[2]}(y)\big)^{\dagger}\check H_{n+N-1}^{-1}\Big]\Omega[n]\begin{bmatrix}\mathcal J_{C_{n-N}^{[1]}}-\mathbf J_{P_{n-N}^{[1]}}\mathcal T\\ \vdots\\ \mathcal J_{C_{n-1}^{[1]}}-\mathbf J_{P_{n-1}^{[1]}}\mathcal T\end{bmatrix}.$$

Hence, we obtain

$$\Big[\big(\check{P}^{[2]}_n(y)\big)^\dagger \check{H}_n^{-1},\dots,\big(\check{P}^{[2]}_{n+N-1}(y)\big)^\dagger \check{H}_{n+N-1}^{-1}\Big]\Omega[n]$$
$$=\Big(W(y)\big(\mathcal{J}_{K^{(pc)}_{n-1}}(y)-\mathbf{J}_{K_{n-1}}(y)\mathcal{T}\big)+\mathcal{J}_{\mathcal{V}}(y)\Big)\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}}-\mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix}^{-1}.$$

Now, for $n\geq N$, from Definition 5.8 and the fact that $\omega_{n,n-N}=\check{H}_n\big(H_{n-N}\big)^{-1}$, we get

$$\big(\check{P}^{[2]}_n(y)\big)^\dagger=\Big(W(y)\big(\mathcal{J}_{K^{(pc)}_{n-1}}(y)-\mathbf{J}_{K_{n-1}}(y)\mathcal{T}\big)+\mathcal{J}_{\mathcal{V}}(y)\Big)\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}}-\mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}}-\mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix}^{-1}\begin{bmatrix}H_{n-N}\\ 0_p\\ \vdots\\ 0_p\end{bmatrix},$$

and the result follows. ∎

Discussion for $n<N$.

Later on, in the context of nonspectral methods, we will derive Corollary 5.39, which can be applied in our monic polynomial perturbation scenario. Thus, we know that $\begin{bmatrix}\mathcal{J}_{C^{[1]}_{0}}-\mathbf{J}_{P^{[1]}_{0}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n}}-\mathbf{J}_{P^{[1]}_{n}}\mathcal{T}\end{bmatrix}$ has full rank.

Proposition 5.21. *The truncations $\omega_{[N]}$, $\check{H}_{[N]}$, and $(\check{S}_2)_{[N]}$ yield the Gauss–Borel factorization*

$$\big(\omega_{[N]}\big)^{-1}\check{H}_{[N]}\big((\check{S}_2)_{[N]}\big)^{-\dagger}=-\begin{bmatrix}\mathcal{J}_{C^{[1]}_{0}}-\mathbf{J}_{P^{[1]}_{0}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{N-1}}-\mathbf{J}_{P^{[1]}_{N-1}}\mathcal{T}\end{bmatrix}\mathcal{R},$$

where $\mathcal{R}$ is given in Lemma 1.40

Proof. From (5.10) we deduce

$$\big(\check{C}^{[1]}(z)\big)_{[N]}W(z)-\big(\check{H}\big(\check{S}_2\big)^{-\dagger}\big)_{[N]}\mathcal{B}(\chi(z))_{[N]}=\omega_{[N]}\big(C^{[1]}(z)\big)_{[N]},$$

so that

$$\begin{bmatrix}\mathcal{J}_{\check{C}^{[1]}_{0}W}\\ \vdots\\ \mathcal{J}_{\check{C}^{[1]}_{N-1}W}\end{bmatrix}-\big(\check{H}\big(\check{S}_2\big)^{-\dagger}\big)_{[N]}\mathcal{B}\mathcal{Q}=\omega_{[N]}\begin{bmatrix}\mathcal{J}_{C^{[1]}_{0}}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{N-1}}\end{bmatrix}.$$

From (5.22) we deduce

$$\begin{bmatrix} \mathcal{J}_{\check{P}_0^{[1]}} \\ \vdots \\ \mathcal{J}_{\check{P}_{N-1}^{[1]}} \end{bmatrix} \mathcal{T} - \left(\check{H}\left(\check{S}_2\right)^{-\dagger}\right)_{[N]} \mathcal{B}\mathcal{Q} = \omega_{[N]} \begin{bmatrix} \mathcal{J}_{C_0^{[1]}} \\ \vdots \\ \mathcal{J}_{C_{N-1}^{[1]}} \end{bmatrix}.$$

Therefore, using (5.7) we conclude

$$-\left(\check{H}\left(\check{S}_2\right)^{-\dagger}\right)_{[N]} = \omega_{[N]} \begin{bmatrix} \mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}} \mathcal{T} \\ \vdots \\ \mathcal{J}_{C_{N-1}^{[1]}} - \mathbf{J}_{P_{N-1}^{[1]}} \mathcal{T} \end{bmatrix} (\mathcal{B}\mathcal{Q})^{-1},$$

and the result follows. ∎

From Proposition 5.21 and using the corresponding explicit quasideterminantal expressions for its solution, see Proposition 1.52 for a similar result with the Hankel block moment matrix, we get

Lemma 5.22. *If $n \in \{0,1,\dots,N-1\}$, then for $0 \leq k < n$ we find the following expressions*

$$\check{H}_n = -\Theta_* \begin{bmatrix} \left(\mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T}\right)\mathcal{R}_n \\ \vdots \\ \left(\mathcal{J}_{C_{n-1}^{[1]}} - \mathbf{J}_{P_{n-1}^{[1]}}\mathcal{T}\right)\mathcal{R}_n \end{bmatrix}, \quad \omega_{n,k} = \Theta_* \left[\begin{array}{c|c} \left(\mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T}\right)\mathcal{R}_n & \\ \vdots & e_k \\ \left(\mathcal{J}_{C_n^{[1]}} - \mathbf{J}_{P_n^{[1]}}\mathcal{T}\right)\mathcal{R}_n & \end{array}\right],$$

$$\left((\check{S}_2)^{\dagger}\right)_{n,k} = \Theta_* \begin{bmatrix} \left(\mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T}\right)\mathcal{R}_{n+1} \\ \vdots \\ \left(\mathcal{J}_{C_{n-1}^{[1]}} - \mathbf{J}_{P_{n-1}^{[1]}}\mathcal{T}\right)\mathcal{R}_{n+1} \\ (e_k)^{\dagger} \end{bmatrix}.$$

Here we have used the matrices $e_k = \left[0_p,\dots,0_p,I_p,0_p,\dots,0_p\right]^{\dagger} \in \mathbb{C}^{(n+1)p\times p}$ with all its $p\times p$ blocks being the zero matrix 0_p, but the k-th block is the identity matrix I_p.

Theorem 5.23 (Spectral Christoffel–Geronimus formulas)**.** *For $n < N$ and monic perturbations of the matrix of functionals we have the following last quasi-determinant expressions for the perturbed bi-orthogonal matrix polynomials*

$$\check{P}_n^{[1]}(x) = \Theta_* \begin{bmatrix} \left(\mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T}\right)\mathcal{R}_n & P_0^{[1]}(x) \\ \vdots & \vdots \\ \left(\mathcal{J}_{C_n^{[1]}} - \mathbf{J}_{P_n^{[1]}}\mathcal{T}\right)\mathcal{R}_n & P_n^{[1]}(x) \end{bmatrix},$$

$$\big(\check{P}^{[2]}_n(x)\big)^\dagger = \Theta_* \begin{bmatrix} \big(\mathcal{J}_{C^{[1]}_0} - \mathbf{J}_{P^{[1]}_0}\mathcal{T}\big)\mathcal{R}_{n+1} \\ \vdots \\ \big(\mathcal{J}_{C^{[1]}_{n-1}} - \mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\big)\mathcal{R}_{n+1} \\ (\chi(x))^\dagger_{[n+1]} \end{bmatrix}.$$

Proof. From (5.7) and (1.8) for $n < N$ we have

$$\check{P}^{[1]}_n(x) = P^{[1]}_n(x) + \big[\omega_{n,0}, \dots, \omega_{n,n-1}\big] \begin{bmatrix} P^{[1]}_0(x) \\ \vdots \\ P^{[1]}_{n-1}(x) \end{bmatrix},$$

$$\big(\check{P}^{[2]}_n\big)^\dagger(x) = I_p x^n + \big[I_p, \dots, I_p x^{n-1}\big] \begin{bmatrix} \big((\check{S}_2)^\dagger\big)_{0,n} \\ \vdots \\ \big((\check{S}_2)^\dagger\big)_{n-1,n} \end{bmatrix}.$$

Then, the result follows from Lemma 5.22. ∎

5.3 Nonspectral Christoffel–Geronimus formulas

We now present an alternative approach on orthogonality relations to the Christoffel–Geronimus formulas which avoids the use of the second kind functions and the spectral structure of the perturbing polynomial. An important feature of these results is that they hold even for perturbing matrix polynomials with singular leading coefficient. The reader will note the similarity between Chapter 3 and this section.

Definition 5.24. *For a given perturbed matrix of functionals $\check{u}$ as in (5.1) we define a semi-infinite block matrix*

$$\begin{aligned} R &:= \left\langle P^{[1]}, \chi(x) \right\rangle_{\check{u}} \\ &= \left\langle P^{[1]}, \chi(x) \right\rangle_{uW^{-1}} + \left\langle P^{[1]}, \chi(x) \right\rangle_{v}, \end{aligned}$$

with block entries

$$R_{n,l} = \left\langle P^{[1]}_n(x), I_p x^l \right\rangle_{\check{u}} \in \mathbb{C}^{p\times p}.$$

Observe that for a Geronimus perturbation of a Borel measure $\mathrm{d}\mu(x)$ we have

$$R_{n,l} = \int P^{[1]}_n(x) x^l \,\mathrm{d}\mu(x) (W(x))^{-1} + \sum_{a=1}^{q} \sum_{j=1}^{s_a} \sum_{m=0}^{\kappa^{(a)}_j - 1} \frac{1}{m!} \Big(P^{[1]}_n(x) x^l \xi^{[a]}_{j,m} l^{(a)}_j(x)\Big)^{(m)}_{x=x_a}.$$

Proposition 5.25. *The following relations hold*

$$R = S_1 \check{M}, \tag{5.25}$$

$$\omega R = \check{H}\big(\check{S}_2\big)^{-\dagger}, \tag{5.26}$$

$$RW(\Lambda^\top) = H\big(S_2\big)^{-\dagger}. \tag{5.27}$$

Proof. (5.25) follows from Definition 5.24. Indeed,

$$\begin{aligned} R &= \left\langle P^{[1]}, \chi(x) \right\rangle_{\check{u}} \\ &= S_1 \left\langle \chi(x), \chi(x) \right\rangle_{\check{u}}. \end{aligned}$$

To deduce (5.26) we recall (5.2), (5.25), and the Gauss factorization of the perturbed matrix of moments. Finally, to get (5.27) we use (5.3) together with (5.25), which implies

$$\omega R \left(W(\Lambda^\top) S_2^\dagger H^{-1} \right) = \big(\check{S}_2\big)^{-\dagger} \left(W(\Lambda^\top) S_2^\dagger H^{-1} \right) = \omega,$$

and as the resolvent ω is unitriangular, i.e. nonsingular, we obtain the result. ∎

From (5.26) it immediately follows that

Proposition 5.26. *The matrix ωR has the following structure*

$$\omega R = \begin{bmatrix} \check{H}_0 & * & * & \cdots \\ 0_p & \check{H}_1 & * & \ddots \\ 0_p & 0_p & \check{H}_2 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Proposition 5.27. *The matrix* $\begin{bmatrix} R_{0,0} & \dots & R_{0,n-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \dots & R_{n-1,n-1} \end{bmatrix}$ *is nonsingular.*

Proof. From (5.25) we conclude for the corresponding truncations that $R_{[n]} = (S_1)_{[n]} \check{M}_{[n]}$ is nonsingular, as we are assuming, to ensure the orthogonality, that $\check{M}_{[n]}$ is nonsingular for every $n \in \mathbb{N} \setminus \{0\}$. ∎

Definition 5.28. *Let us introduce the polynomials $r^K_{n,l}(y) \in \mathbb{C}^{p\times p}[y]$, $l \in \{0, \dots, n-1\}$,*

$$\begin{aligned} r^K_{n,l}(y) :&= \left\langle W(y) K_{n-1}(x,y), I_p x^l \right\rangle_{\check{u}} - I_p y^l \\ &= \left\langle W(y) K_{n-1}(x,y), I_p x^l \right\rangle_{uW^{-1}} + \left\langle W(y) K_{n-1}(x,y), I_p x^l \right\rangle_{v} - I_p y^l. \end{aligned}$$

Proposition 5.29. *For $l \in \{0,1,\dots,n-1\}$ and $m=\min(n,N)$,*

$$r^K_{n,l} = \Big[(\check{P}^{[2]}_n(y))^\dagger(\check{H}_n)^{-1},\dots,(\check{P}^{[2]}_{n-1+N}(y))^\dagger(\check{H}_{n-1+N})^{-1}\Big]\Omega[n]\begin{bmatrix}R_{n-m,l}\\ \vdots\\ R_{n-1,l}\end{bmatrix}.$$

Proof. It follows from (5.12), Definition 5.24, and (1.15). ■

Definition 5.30. *For $n\geq N$, given the matrix*

$$\begin{bmatrix}R_{n-N,0} & \dots & R_{n-N,n-1}\\ \vdots & & \vdots\\ R_{n-1,0} & \dots & R_{n-1,n-1}\end{bmatrix}\in\mathbb{R}^{Np\times np},$$

we construct a submatrix labelled by selecting Np columns among all the np columns. For that aim, we use indexes (i,a) labeling the columns, where i runs through $\{0,\dots,n-1\}$ and indicates the block, and $a\in\{1,\dots,p\}$ denotes the corresponding column in that block, i.e. (i,a) is an index selecting the a-th column of the i-block. Given a set of Np different pairs $I=\{(i_r,a_r)\}_{r=1}^{Np}$, (recall that $i_r\in\{0,\cdots,n-1\}$ and $a_r\in\{1,\cdots,p\}$) with a lexicographic ordering, we define the corresponding submatrix $R^\square_n := \big[\mathfrak{c}_{(i_1,a_1)},\dots,\mathfrak{c}_{(i_{Np},a_{Np})}\big]$. Here $\mathfrak{c}_{(i_r,a_r)}$ denotes the a_r-th column of the matrix

$$\begin{bmatrix}R_{n-N,i_r}\\ \vdots\\ R_{n-1,i_r}\end{bmatrix}.$$

The set of indexes I is said to be poised if $R^\square_n$ is nonsingular. We also use the notation $r^\square_n := \big[\tilde{\mathfrak{c}}_{(i_1,a_1)},\dots,\tilde{\mathfrak{c}}_{(i_{Np},a_{Np})}\big]$. Here $\tilde{\mathfrak{c}}_{(i_r,a_r)}$ denotes the a_r-th column of the matrix R_{n,i_r}. Given a poised set on indexes we define $(r^K_n(y))^\square$ as the matrix built up by taking from the matrices $r^K_{n,i_r}(y)$ the columns a_r.

Lemma 5.31. *For $n\geq N$, there exists at least a poised set.*

Proof. For $n\geq N$, let us consider the rectangular block matrix

$$\begin{bmatrix}R_{n-N,0} & \dots & R_{n-N,n-1}\\ \vdots & & \vdots\\ R_{n-1,0} & \dots & R_{n-1,n-1}\end{bmatrix}\in\mathbb{C}^{Np\times np}.$$

As the truncation $R_{[n]}$ is nonsingular, all its Np rows are linearly independent. Thus, there must be Np linearly independent columns and the desired result follows. ■

Lemma 5.32. *If the leading coefficient A_N of the perturbing polynomial $W(x)$ is nonsingular, then we can decompose any monomial $I_p x^l$ in the following way*

$$I_p x^l = \alpha_l(x) W(x)^\dagger + \beta_l(x),$$

where $\alpha_l(x), \beta_l(x) = \beta_{l,0} + \cdots + \beta_{l,N-1} x^{N-1} \in \mathbb{R}^{p\times p}[x]$, *with* $\deg \alpha_l(x) = l - N$.

Proposition 5.33. *Let us assume that the matrix polynomial $W(x) = A_N x^N + \cdots + A_0$ has a nonsingular leading coefficient and $n \geq N$. Then, the matrix*

$$\begin{bmatrix} R_{n-N,0} & & R_{n-N,N-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \dots & R_{n-1,N-1} \end{bmatrix}$$

is nonsingular.

Proof. From Proposition 5.26 or, equivalently, (5.7) we deduce

$$\begin{bmatrix} \omega_{n,n-N}, \dots, \omega_{n,n-1} \end{bmatrix} \begin{bmatrix} R_{n-N,l} \\ \vdots \\ R_{n-1,l} \end{bmatrix} = -R_{n,l},$$

for $l \in \{0,1,\dots,n-1\}$. In particular, the resolvent vector $\begin{bmatrix} \omega_{n,n-N}, \dots, \omega_{n,n-1} \end{bmatrix}$ is a solution of the linear system

$$\begin{bmatrix} \omega_{n,n-N}, \dots, \omega_{n,n-1} \end{bmatrix} \begin{bmatrix} R_{n-N,0} & & R_{n-N,N-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \dots & R_{n-1,N-1} \end{bmatrix} = -\begin{bmatrix} R_{n,0}, \dots, R_{n,N-1} \end{bmatrix}. \tag{5.28}$$

We will show now that this is the unique solution of the system. Let us proceed by contradiction. Let us assume that there is another solution $\begin{bmatrix} \tilde{\omega}_{n,n-N}, \dots, \tilde{\omega}_{n,n-1} \end{bmatrix}$. Consider the monic matrix polynomial

$$\tilde{P}_n(x) = P_n^{[1]}(x) + \tilde{\omega}_{n,n-1} P_{n-1}^{[1]}(x) + \cdots + \tilde{\omega}_{n,n-N} P_{n-N}^{[1]}(x).$$

Because $\begin{bmatrix} \tilde{\omega}_{n,n-N}, \dots, \tilde{\omega}_{n,n-1} \end{bmatrix}$ solves (5.28) we know that

$$\langle \tilde{P}_n(x), I_p x^l \rangle_{\check{u}} = 0_p, \qquad l \in \{0, \dots, N-1\}.$$

Lemma 5.32 implies the following relations for $\deg \alpha_l(x) < m$,

$$\begin{aligned} \langle P_m^{[1]}(x), I_p x^l \rangle_{\check{u}} &= \langle P_m^{[1]}(x), \alpha_l(x) \rangle_{\check{u}W} + \langle P_m^{[1]}(x), \beta_l(x) \rangle_{\check{u}} \\ &= \langle P_m^{[1]}(x), \alpha_l(x) \rangle_{u} + \langle P_m^{[1]}(x), \beta_l(x) \rangle_{\check{u}} \end{aligned}$$

$$= \langle P_m^{[1]}(x), \beta_l(x) \rangle_{\check{u}}.$$

But $\deg \alpha_l(x) \leq l - N$, so the previous equation will hold at least for $l - N < m$, i.e. $l < m + N$. Consequently, for $l \in \{0, \dots, n-1\}$, we find

$$\begin{aligned}
\langle \tilde{P}_n(x), x^l I_p \rangle_{\check{u}} &= \sum_{k=n-N}^{n} \tilde{\omega}_{n,k} \langle P_k^{[1]}(x), \alpha_l(x) W(x)^\dagger \rangle_{\check{u}} + \langle \tilde{P}_n(x), \beta_l(x) \rangle_{\check{u}} \\
&= \sum_{k=n-N}^{n} \tilde{\omega}_{n,k} \langle P_k^{[1]}(x), \alpha_l(x) \rangle_u + \langle \tilde{P}_n(x), \beta_l(x) \rangle_{\check{u}} \\
&= 0_p.
\end{aligned}$$

Therefore, from the uniqueness of the bi-orthogonal families, we deduce

$$\tilde{P}_n(x) = \check{P}_n^{[1]}(x),$$

and, recalling (5.7), there is an unique solution of (5.28). Thus,

$$\begin{bmatrix} R_{n-N,0} & & R_{n-N,n-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \dots & R_{n-1,n-1} \end{bmatrix}$$

is nonsingular. ∎

Proposition 5.34. *For $n < N$, we can write*

$$\begin{bmatrix} \omega_{n,0}, \dots, \omega_{n,n-1} \end{bmatrix} = - \begin{bmatrix} R_{n,0}, \dots, R_{n,n-1} \end{bmatrix} \begin{bmatrix} R_{0,0} & \dots & R_{0,n-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \dots & R_{n-1,n-1} \end{bmatrix}^{-1}$$

and, for $n \geq N$ and a given poised set I, which always exists, we have

$$\begin{bmatrix} \omega_{n,n-N}, \dots, \omega_{n,n-1} \end{bmatrix} = - r_n^{\square} (R_n^{\square})^{-1}.$$

Proof. It is a straightforward consequence of Lemma 5.31. ∎

Theorem 5.35 (Non-spectral Christoffel-Geronimus formulas)**.** *Given a matrix Geronimus transformation the corresponding perturbed monic bi-orthogonal polynomials, $(\check{P}_n^{[1]}(x), \check{P}_n^{[2]}(x))_{n \in \mathbb{N}}$, can be expressed as follows. For $n \geq N$,*

$$\check{P}_n^{[1]}(x) = \Theta_* \left[\begin{array}{c|c} R_n^{\square} & \begin{matrix} P_{n-N}^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x) \end{matrix} \\ \hline r_n^{\square} & P_n^{[1]}(x) \end{array} \right], \qquad \check{H}_n = \Theta_* \left[\begin{array}{c|c} R_n^{\square} & \begin{matrix} R_{n-N,n} \\ \vdots \\ R_{n-1,n} \end{matrix} \\ \hline r_n^{\square} & R_{n,n} \end{array} \right],$$

$$
\left(\check{P}_n^{[2]}(y)\right)^{\dagger} A_N = -\Theta_* \left[\begin{array}{c|c} \multirow{4}{*}{$R_n^{\square}$} & H_{n-N} \\ & 0_p \\ & \vdots \\ & 0_p \\ \hline (r_n^K(y))^{\square} & 0_p \end{array}\right].
$$

For $n < N$, we have $\check{H}_n = \Theta_*\left(R_{[n+1]}\right)$ and

$$
\check{P}_n^{[1]}(x) = \Theta_* \begin{bmatrix} R_{0,0} & \dots & R_{0,n-1} & P_0^{[1]}(x) \\ \vdots & & \vdots & \vdots \\ R_{n,0} & \dots & R_{n,n-1} & P_n^{[1]}(x) \end{bmatrix}, \quad \left(\check{P}_n^{[2]}(x)\right)^{\dagger} = \Theta_* \begin{bmatrix} R_{0,0} & \dots & R_{0,n-1} & R_{0,n} \\ \vdots & & \vdots & \vdots \\ R_{n-1,0} & \dots & R_{n-1,n-1} & R_{n-1,n} \\ I_p & \dots & I_p x^{n-1} & I_p x^n \end{bmatrix}.
$$

Proof. From the connection formula (5.7), for $m = \min(n, N)$ we have

$$
\check{P}_n^{[1]}(x) = [\omega_{n,n-m}, \dots, \omega_{n,n-1}] \begin{bmatrix} P_{n-m}^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x) \end{bmatrix} + P_n^{[1]}(x),
$$

and from here

$$
\check{H}_n = [\omega_{n,n-m}, \dots, \omega_{n,n-1}] \begin{bmatrix} R_{n-m,n} \\ \vdots \\ R_{n-1,n} \end{bmatrix} + R_{n,n}.
$$

Then, recalling Proposition 5.34 we obtain the desired results.

For $n \geq N$, we have

$$
r_{n,l}^K(y) = \left[(\check{P}_n^{[2]}(y))^{\dagger}(\check{H}_n)^{-1}, \dots, (\check{P}_{n-1+N}^{[2]}(y))^{\dagger}(\check{H}_{n-1+N})^{-1}\right] \Omega[n] \begin{bmatrix} R_{n-N,l} \\ \vdots \\ R_{n-1,l} \end{bmatrix}.
$$

Then, we get

$$
(r_n^K(y))^{\square}(R_n^{\square})^{-1} = \left[(\check{P}_n^{[2]}(y))^{\dagger}(\check{H}_n)^{-1}, \dots, (\check{P}_{n-1+N}^{[2]}(y))^{\dagger}(\check{H}_{n-1+N})^{-1}\right] \Omega[n].
$$

In particular, recalling (5.6) we deduce that

$$
(\check{P}_n^{[2]}(y))^{\dagger} A_N = (r_n^K(y))^{\square}(R_n^{\square})^{-1} \begin{bmatrix} H_{n-N} \\ 0_p \\ \vdots \\ 0_p \end{bmatrix}.
$$

For $n<N$, we can write (5.26) as $R\big(\check{S}_2\big)^{\dagger}=\omega^{-1}\check{H}$, which implies

$$\begin{bmatrix}\big((S_2)^{\dagger}\big)_{0,n}\\ \vdots\\ \big((S_2)^{\dagger}\big)_{n-1,n}\end{bmatrix}=-\big(R_{[n]}\big)^{-1}\begin{bmatrix}R_{0,n}\\ \vdots\\ R_{n-1,n}\end{bmatrix}.$$

Therefore,

$$\begin{aligned}\big(P^{[2]}_n(x)\big)^{\dagger}&=x^n+\big[I_p,\dots,I_px^{n-1}\big]\begin{bmatrix}\big((S_2)^{\dagger}\big)_{0,n}\\ \vdots\\ \big((S_2)^{\dagger}\big)_{n-1,n}\end{bmatrix}\\&=x^n-\big[I_p,\dots,I_px^{n-1}\big]\big(R_{[n]}\big)^{-1}\begin{bmatrix}R_{0,n}\\ \vdots\\ R_{n-1,n}\end{bmatrix}.\end{aligned}$$

■

Observe the Gauss–Borel factorization

$$R_{[n]}=\big(\omega_{[n]}\big)^{-1}\check{H}_{[n]}\big((\check{S}_2)_{[n]}\big)^{-\dagger},$$

for $n<N$ gives the corresponding formulas in the previous theorem.

5.4 Spectral versus nonspectral

Definition 5.36. *We introduce the truncation obtained by taking only the first N columns of a given semi-infinite matrix*

$$R^{(N)}:=\begin{bmatrix}R_{0,0} & R_{0,1} & \dots R_{0,N-1}\\ R_{1,0} & R_{0,1} & \dots R_{1,N-1}\\ \vdots & & \vdots\end{bmatrix}.$$

Then, we can connect the spectral and nonspectral techniques as follows.

Proposition 5.37. *The following relation holds*

$$\mathcal{J}_{C^{[1]}}-\mathbf{J}_{P^{[1]}}\mathcal{T}_G=-R^{(N)}\mathcal{B}\mathcal{Q}.$$

Proof. From (5.10) we deduce that

$$\check{C}^{[1]}(x)W(x)-\check{H}\big(\check{S}_2\big)^{-\dagger}\begin{bmatrix}\mathcal{B}(\chi(x))_{[N]}\\ 0_p\\ \vdots\end{bmatrix}=\omega C^{[1]}(x).$$

Therefore,

$$\mathcal{J}_{\check{C}^{[1]}W} - \check{H}\big(\check{S}_2\big)^{-\dagger}\begin{bmatrix}\mathcal{B}\mathcal{Q}\\ 0_p\\ \vdots\end{bmatrix} = \omega\mathcal{J}_{C^{[1]}},$$

and recalling (5.22) and (5.23), we get

$$\omega\big(\mathcal{J}_{C^{[1]}} - \mathbf{J}_{P^{[1]}}\mathcal{T}\big) = -\check{H}\big(\check{S}_2\big)^{-\dagger}\begin{bmatrix}\mathcal{B}\mathcal{Q}\\ 0_p\\ \vdots\end{bmatrix}.$$

Now, from (5.26) and since ω is unitriangular the result follows. ■

We now discuss an important fact, which ensures that the spectral Christoffel–Geronimus formulas make sense

Corollary 5.38. *When the leading coefficient A_N is nonsingular and $n \geq N$,*

$$\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}} - \mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}} - \mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix}$$

is nonsingular.

Proof. From Proposition 5.37 one deduces the following formula

$$\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}} - \mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n-1}} - \mathbf{J}_{P^{[1]}_{n-1}}\mathcal{T}\end{bmatrix} = -\begin{bmatrix}R_{n-N,0} & & R_{n-N,N-1}\\ \vdots & & \vdots\\ R_{n-1,0} & \dots & R_{n-1,N-1}\end{bmatrix}\mathcal{B}\mathcal{Q}. \tag{5.29}$$

Now, recalling Proposition 5.33 and Lemma 1.40 we get the result. ■

We stress at this point that (5.29) connects the spectral and nonspectral methods. Moreover, when we border with a further block row we also have

$$\begin{bmatrix}\mathcal{J}_{C^{[1]}_{n-N}} - \mathbf{J}_{P^{[1]}_{n-N}}\mathcal{T}\\ \vdots\\ \mathcal{J}_{C^{[1]}_{n}} - \mathbf{J}_{P^{[1]}_{n}}\mathcal{T}\end{bmatrix} = -\begin{bmatrix}R_{n-N,0} & & R_{n-N,N-1}\\ \vdots & & \vdots\\ R_{n,0} & \dots & R_{n,N-1}\end{bmatrix}\mathcal{B}\mathcal{Q}.$$

Corollary 5.39. *When the leading coefficient A_N is nonsingular and $n<N$, then*

$$\begin{bmatrix} \mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T} \\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}} - \mathbf{J}_{P_{n-1}^{[1]}}\mathcal{T} \end{bmatrix} \mathcal{R}_n$$

is nonsingular and the matrix

$$\begin{bmatrix} \mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T} \\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}} - \mathbf{J}_{P_{n-1}^{[1]}}\mathcal{T} \end{bmatrix}$$

has full rank.

Proof. From Proposition 5.37 we know that when the leading coefficient A_N is nonsingular the following relation holds

$$\left(\mathcal{J}_{C^{[1]}} - \mathbf{J}_{P^{[1]}}\mathcal{T}\right)\mathcal{R} = -R^{(N)},$$

where $\mathcal{R} = (Y, JY, \dots, J^{N-1}Y)$ is constructed in terms of a canonical Jordan triple (X,Y,J) of the perturbing polynomial $W(x)$ (see Lemma 1.40). Thus, for $n<N$ we have in terms of the truncation $\mathcal{R}_n = (Y, JY, \dots, J^{n-1}Y)$, which as we have seen is a full rank matrix,

$$\begin{bmatrix} \mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T} \\ \vdots \\ \mathcal{J}_{C_{n-1}^{[1]}} - \mathbf{J}_{P_{n-1}^{[1]}}\mathcal{T} \end{bmatrix} \mathcal{R}_n = - \begin{bmatrix} R_{0,0} & & R_{0,n-1} \\ \vdots & & \vdots \\ R_{n,0} & \dots & R_{n,n-1} \end{bmatrix}.$$

Thus, as the matrix in the right hand side is nonsingular, we deduce the result. ■

5.40 Example . *Let $(L_n^{(4)}(x))_{n\in\mathbb{N}}$ be the sequence of monic Laguerre polynomials which are orthogonal with respect to the measure $d\mu = x^4 e^{-x}dx$, supported on $(0,\infty)$. On the other hand, let*

$$W(x) = \begin{bmatrix} x^2 & 1 \\ 0 & x^2 \end{bmatrix} \in \mathbb{C}^{2\times 2}[x]. \quad \textit{with} \quad W^{-1}(x) = \frac{1}{x^4}\begin{bmatrix} x^2 & -1 \\ 0 & x^2 \end{bmatrix}.$$

$x_0 = 0$ is the only zero of $W(x)$ (with multiplicity 4) and thus its associated Jordan chain of length 4 is $\{r_{1,0}^{(1)}, r_{1,1}^{(1)}, r_{1,2}^{(1)}, r_{1,3}^{(1)}\}$, where $r_{1,0}^{(1)} = r_{1,1}^{(1)} = (1,0)^\top$ and $r_{1,2}^{(1)} = r_{1,3}^{(1)} = (0,-1)^\top$. Moreover, the left root polynomial is $l_1^{(1)}(x) = (0,1) + (0,1)x - (1,0)x^2 - (1,0)x^3$.
Let $d\mu = x^4 e^{-x} I_2 dx$ be a matrix of measures. Then the most general sesquilinear form $\langle\cdot,\cdot\rangle_{\check{\mu}}$ such that $\left\langle P(x), Q(x)W(x)^\dagger\right\rangle_{\check{\mu}} = \langle P(x), Q(x)\rangle_\mu$ is

$$\langle P, Q\rangle_{\check{\mu}} = \int P(x)\,\mathrm{d}\mu(x)(W(x))^{-1}Q(x)^\dagger + \sum_{m=0}^{3}\frac{1}{m!}\left(P(x)\xi_m l_1^{(a)}(x)Q(x)^\dagger\right)_{x=0}^{(m)}. \tag{5.30}$$

In particular, if we take $\xi_0 = \xi_1 = (1,1)^\dagger$ *and* $\xi_2 = \xi_3 = (0,0)^\dagger$, *then* (5.30) *becomes*

$$\langle P, Q\rangle_{\check{\mu}} = \int_0^\infty P(x)\begin{pmatrix} x^2 & -1 \\ 0 & x^2 \end{pmatrix} Q^\top(x) e^{-x}dx + \begin{bmatrix} P(0) & P'(0) \end{bmatrix} \begin{bmatrix} 0&1&0&1 \\ 0&1&0&1 \\ 0&1&0&0 \\ 0&1&0&0 \end{bmatrix} \begin{bmatrix} Q(0)^\dagger \\ (Q')(0)^\dagger \end{bmatrix}.$$

If we use the property i) of Proposition 4.21, we can compute

$$R_{n,0} = \left\langle L_n^4(x)I_2, I_2\right\rangle_{\check{\mu}} = \begin{bmatrix} 2(-1)^n(n+1)! & \frac{(-1)^n}{24}(1-\frac{n}{5})(n+4)! + \frac{(-1)^{n+1}}{6}(n+3)! \\ 0 & \frac{(-1)^n}{24}(1-\frac{n}{5})(n+4)! + 2(-1)^n(n+1)! \end{bmatrix}$$

$$R_{n,1} = \left\langle L_n^4(x)I_2, xI_2\right\rangle_{\check{\mu}} = \begin{bmatrix} 6(-1)^n n! & \frac{(-1)^{n+1}}{2}(n-2)! + \frac{(-1)^n}{24}(n+4)! \\ 0 & 6(-1)^n n! + \frac{(-1)^n}{24}(n+4)! \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} R_{n-2,0} & R_{n-2,1} \\ R_{n-1,0} & R_{n-1,1} \end{bmatrix} =$$

$$\begin{bmatrix} 2(-1)^n(n-1)! & \frac{(-1)^n}{24}(1-\frac{n-2}{5})(n+2)! + \frac{(-1)^{n+1}}{6}(n+1)! & 6(-1)^n(n-2)! & \frac{(-1)^{n+1}}{2}(n-4)! + \frac{(-1)^n}{24}(n+2)! \\ 0 & \frac{(-1)^n}{24}(1-\frac{n-2}{5})(n+2)! + 2(-1)^n(n-1)! & 0 & 6(-1)^n(n-2)! + \frac{(-1)^n}{24}(n+2)! \\ 2(-1)^{n-1}n! & \frac{(-1)^{n-1}}{24}(1-\frac{n-1}{5})(n+3)! + \frac{(-1)^n}{6}(n+2)! & 6(-1)^{n-1}(n-1)! & \frac{(-1)^n}{2}(n-3)! + \frac{(-1)^{n-1}}{24}(n+3) \\ 0 & \frac{(-1)^{n-1}}{24}(1-\frac{n-1}{5})(n+3)! + 2(-1)^{n-1}n! & 0 & 6(-1)^{n-1}(n-1)! + \frac{(-1)^{n-1}}{24}(n+4)! \end{bmatrix}.$$

On the other hand, from property vi) in Proposition 4.21 we have

$$K_n^\alpha(x,y) = \frac{K_{n+m}^{\alpha-m}(x,y)}{y^m} - \sum_{k=0}^{m-1} \frac{L_{n+m+1}^{\alpha-k}(x)}{\Gamma(\alpha+n+m-k+1)(n+m+1)!} \frac{L_{n+m+1}^{\alpha-k-1}(y)}{y^{k+1}}.$$

Thus, if α *is a nonnegative integer number such that* $0 \le \alpha - m \le n+m$, *then*

$$\int K_n^\alpha(x,y) x^{\alpha-m} e^{-x}dx = \frac{1}{y^m} - \sum_{k=0}^{m-1} \frac{(-1)^{n+m+1}(m-k+n+m)!\Gamma(\alpha-m+1)}{\Gamma(\alpha+n+m-k+1)(n+m+1)!(m-k-1)!} \frac{L_{n+m+1}^{\alpha+k-1}(y)}{y^{k+1}}.$$

Observe that the above allows us to give an explicit formula for the matrix $r_n^K(y)$. *Thus the sequences of matrix bi-orthogonal polynomials* $(\check{L}_n^{[1]}(x), \check{L}_n^{[2]}(x))_{n\in\mathbb{N}}$ *with respect to* $\langle\cdot,\cdot\rangle_{\check{\mu}}$ *are*

$$\check{L}_n^{[1]}(x) = \Theta_* \left[\begin{array}{cc|c} R_{n-2,0} & R_{n-2,1} & L_{n-2}^4(x)I_2 \\ R_{n-1,0} & R_{n-1,1} & L_{n-1}^4(x)I_2 \\ \hline R_{n,0} & R_{n,1} & L_n^4(x)I_2 \end{array}\right],$$

$$\left(\check{L}_n^{[2]}(y)\right)^\dagger = -\Theta_* \left[\begin{array}{cc|c} R_{n-2,0} & R_{n-2,1} & (n-2)!\Gamma(n+3)I_2 \\ R_{n-1,0} & R_{n-1,1} & 0_2 \\ \hline \multicolumn{2}{c|}{r_n^K(y)} & 0_2 \end{array}\right].$$

5.5 Applications

5.5.1 Unimodular perturbations and nonspectral techniques

The spectral methods apply to those Geronimus transformations with a perturbing polynomial $W(x)$ having a nonsingular leading coefficient A_N. This was also the case where the techniques developed in 4 for matrix Christoffel transformations are considered assuming the perturbing polynomial has a nonsingular leading coefficient. However, we have shown that despite we can extend the use of the spectral techniques to the study of matrix Geronimus transformations, we also have a nonspectral approach applicable even for singular leading coefficients. For example, some cases that have appeared several times in the literature are unimodular perturbations and, consequently, $W(x)$ has a singular leading coefficient. In this case, we have that $(W(x))^{-1}$ is a matrix polynomial and we can consider the Geronimus transformation associated with the matrix polynomial $(W(x))^{-1}$ which can be understood as a Christoffel transformation with perturbing matrix polynomial $W(x)$. Since $\sigma(W)=\varnothing$, the perturbation has no masses and it can be understood as a Christoffel transformation

$$\check{u}=u\big((W(x))^{-1}\big)^{-1}=uW(x).$$

We can apply Theorem 5.35 with

$$R=\left\langle P^{[1]},\chi(x)\right\rangle_{uW}, \qquad R_{n,l}=\left\langle P^{[1]}_n,I_px^l\right\rangle_{uW}\in\mathbb{C}^{p\times p}.$$

If the matrix of linear functionals is just a matrix of measures μ supported on the real line, then we can write

$$R_{n,l}=\int P^{[1]}_n(x)x^l\,\mathrm{d}\mu(x)W(x).$$

5.41 Example . *If we deal with a scalar case, so that $u=u_0I_p$ where $u_0\in(\mathbb{C}[x])'$ is a scalar linear functional and $W(x)=(W(x))^{\dagger}$, then the sequences $(\check{P}^{[1]}_n(x))_{n\in\mathbb{N}}$ and $(\check{P}^{[2]}_n(x))_{n\in\mathbb{N}}$ coincide and $\check{H}_n=(\check{H}_n)^{\dagger}$ is Hermitian. If $(p_n(x))_{n\in\mathbb{N}}$ denotes the set of orthogonal polynomials on the real line with respect to u_0, we have*

$$R_{n,l}=\big\langle u_0,p_n(x)x^lW(x)\big\rangle.$$

For example, if we take $p=2$ and the unimodular perturbation given by

$$W(x)=\begin{bmatrix}(A_2)_{1,1}x^2+(A_1)_{1,1}x+(A_0)_{1,1} & (A_2)_{1,2}x^2+(A_1)_{1,2}x+(A_0)_{1,2}\\ (A_2)_{1,2}x^2+(A_1)_{1,2}x+(A_0)_{1,2} & (A_2)_{2,2}x^2+(A_1)_{2,2}x+(A_0)_{2,2}\end{bmatrix}$$

we have that its inverse is the matrix polynomial

$$(W(x))^{-1}=\frac{1}{\det W(x)}\begin{bmatrix}(A_2)_{2,2}x^2+(A_1)_{2,2}x+(A_0)_{2,2} & -(A_2)_{1,2}x^2-(A_1)_{1,2}x-(A_0)_{1,2}\\ -(A_2)_{1,2}x^2-(A_1)_{1,2}x-(A_0)_{1,2} & (A_2)_{1,1}x^2+(A_1)_{1,1}x+(A_0)_{1,1}\end{bmatrix},$$

and since $\det W(x)$ *is a constant, the inverse has also degree 2. Therefore, for* $n \in \{2,3,\dots\}$*, we get the following expressions for the perturbed matrix orthogonal polynomials*

$$\check{P}_n(x) = \Theta_* \left[\begin{array}{c|c} R_n^{\square} & \begin{matrix} p_{n-2}(x)I_2 \\ p_{n-1}(x)I_2 \end{matrix} \\ \hline r_n^{\square} & p_n(x)I_2 \end{array}\right], \quad \check{H}_n = \Theta_* \left[\begin{array}{c|c} R_n^{\square} & \begin{matrix} \langle u_0, p_{n-2}(x)x^n W(x)\rangle \\ \langle u_0, p_{n-1}(x)x^n W(x)\rangle \end{matrix} \\ \hline r_n^{\square} & \langle u_0, p_n(x)x^n W(x)\rangle \end{array}\right],$$

where $R_n^{\square}$ *is a square matrix of size* 4×4*. Moreover, since* $R_{n,l} = 0_{2\times 2}$ *for* $l = 0,\dots,n-3$*, then its columns are taken from the following matrix*

$$\begin{bmatrix} \langle u_0, p_{n-2}(x)x^{n-4}W(x)\rangle & \langle u_0, p_{n-2}(x)x^{n-3}W(x)\rangle & \langle u_0, p_{n-2}(x)x^{n-2}W(x)\rangle & \langle u_0, p_{n-2}(x)x^{n-1}W(x)\rangle \\ 0_{2\times 2} & \langle u_0, p_{n-1}(x)x^{n-3}W(x)\rangle & \langle u_0, p_{n-1}(x)x^{n-2}W(x)\rangle & \langle u_0, p_{n-1}(x)x^{n-1}W(x)\rangle \end{bmatrix}.$$

5.42 Example . *If our perturbation is* $W(x) = I_p + p(x)E_{i,j}$*, where the matrix* $E_{i,j}$ *was defined in Subsection 4.4.1 and* $p(x) \in \mathbb{C}[x]$ *with* $\deg p(x) = N$*, then we have that* $(W(x))^{-1} = I_p - p(x)E_{i,j}$ *and* $\deg W = \deg(W^{-1}) = N$*. If we assume an initial matrix of linear functionals u, for* $n \geq N$*, the first family of perturbed polynomials will be*

$$\check{P}_n^{[1]}(x) = \Theta_* \begin{bmatrix} \left\langle P_{n-N}^{[1]}(x), x^{k_1}(I_p + p(x)E_{i,j})\right\rangle_u & \dots & \left\langle P_{n-N}^{[1]}(x), x^{k_N}(I_p + p(x)E_{i,j})\right\rangle_u & P_{n-N}^{[1]}(x) \\ \vdots & & \vdots & \vdots \\ \left\langle P_n^{[1]}(x), x^{k_1}(I_p + p(x)E_{i,j})\right\rangle_u & \dots & \left\langle P_n^{[1]}(x), x^{k_N}(I_p + p(x)E_{i,j})\right\rangle_u & P_n^{[1]}(x) \end{bmatrix}.$$

Here, the sequence of different integers $\{k_1,\dots,k_N\} \subset \{1,\dots,n-1\}$ *satisfies*

$$\begin{vmatrix} \left\langle P_{n-N}^{[1]}(x), x^{k_1}(I_p + p(x)E_{i,j})\right\rangle_u & \dots & \left\langle P_{n-N}^{[1]}(x), x^{k_N}(I_p + p(x)E_{i,j})\right\rangle_u \\ \vdots & & \vdots \\ \left\langle P_{n-1}^{[1]}(x), x^{k_1}(I_p + p(x)E_{i,j})\right\rangle_u & \dots & \left\langle P_{n-1}^{[1]}(x), x^{k_N}(I_p + p(x)E_{i,j})\right\rangle_u \end{vmatrix} \neq 0.$$

Observe that in the two previous cases unimodular matrix polynomials are particularly simple, since the degree of the perturbation and its inverse coincide. However, this is not always true. A simple example illustrates this situation.

5.43 Example . *If now we take the perturbation* $W(x) = \begin{bmatrix} 1 & x^3 & x^2 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{bmatrix}$*, then* $(W(x))^{-1} = \begin{bmatrix} 1 & -x^3 & x^4 - x^2 \\ 0 & 1 & -x \\ 0 & 0 & 1 \end{bmatrix}$*. Notice that* $\deg(W(x)) = 3$ *and* $\deg(W(x))^{-1} = 4$*. Thus*

$$\check{P}_n^{[1]}(x) = \Theta_* \left[\begin{array}{ccc|c} \left\langle P_{n-4}^{[1]}(x), x^{k_1}W(x)\right\rangle_u & \dots & \left\langle P_{n-4}^{[1]}(x), x^{k_4}W(x)\right\rangle_u & P_{n-4}^{[1]}(x) \\ \vdots & & \vdots & \vdots \\ \left\langle P_{n-1}^{[1]}(x), x^{k_1}W(x)\right\rangle_u & \dots & \left\langle P_{n-1}^{[1]}(x), x^{k_4}W(x)\right\rangle_u & P_{n-1}^{[1]}(x) \\ \hline \left\langle P_n^{[1]}(x), x^{k_1}W(x)\right\rangle_u & \dots & \left\langle P_n^{[1]}(x), x^{k_4}W(x)\right\rangle_u & P_n^{[1]}(x) \end{array}\right],$$

$$
(\check{P}_n^{[2]}(y))^\dagger \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \Theta_* \left[\begin{array}{ccc|c} \left\langle P^{[1]}_{n-4}(x), x^{k_1} W(x)\right\rangle_u & \dots & \left\langle P^{[1]}_{n-4}(x), x^{k_4} W(x)\right\rangle_u & H_{n-4} \\ \vdots & & \vdots & \vdots \\ \left\langle P^{[1]}_{n}(x), x^{k_1} W(x)\right\rangle_u & \dots & \left\langle P^{[1]}_{n}(x), x^{k_4} W(x)\right\rangle_u & 0_p \\ \hline & (r^K_n(y))^\square & & 0_p \end{array}\right].
$$

Again, in this case the set of different integers $\{k_1,\dots,k_4\} \subset \{1,\dots,n-1\}$ *is such that*

$$
\begin{vmatrix} \left\langle P^{[1]}_{n-4}(x), x^{k_1} W(x)\right\rangle_u & \dots & \left\langle P^{[1]}_{n-4}(x), x^{k_4} W(x)\right\rangle_u \\ \vdots & & \vdots \\ \left\langle P^{[1]}_{n-1}(x), x^{k_1} W(x)\right\rangle_u & \dots & \left\langle P^{[1]}_{n-1}(x), x^{k_4} W(x)\right\rangle_u \end{vmatrix} \neq 0.
$$

5.5.2 Degree one matrix Geronimus transformations

We consider a perturbing polynomial of degree one

$$
W(x) = xI_p - A,
$$

and assume, for the sake of simplicity, that all $\xi^{[a]}_{j,m}$ in (5.1) are zero, i.e., there are no masses. Observe that in this case a Jordan pair (X,J) is such that $A = XJX^{-1}$, and Lemma 1.37 implies that the root spectral jet of a polynomial $P(x) = \sum_k P_k x^k \in \mathbb{C}^{p\times p}[x]$ is $\mathcal{J}_P = P(A)X$, where we understand a right evaluation, i.e. $P(A) := \sum_k P_k A^k$. For $\sigma(A) \cap \operatorname{supp}(u) = \varnothing$, a similar argument yields

$$
\mathcal{J}_{C_n^{[1]}} = \left\langle P^{[1]}(x), (A^\dagger - I_p x)^{-1}\right\rangle X = \left\langle P^{[1]}(x), (A - I_p x)^{-\dagger}\right\rangle X,
$$

expressed in terms of the resolvent $(A - I_p x)^{-1}$ of A. Formally, it can be written

$$
\mathcal{J}_{C_n^{[1]}} = C_n^{[1]}(A)X,
$$

where we again understand a right evaluation in the Taylor series of the Cauchy transform. Moreover, we will also need the root spectral jet of the mixed Christoffel–Darboux kernel

$$
\begin{aligned}
\mathcal{J}_{K^{(pc)}_{n-1}}(y) &= \sum_{k=0}^{n-1} \big(P^{[2]}_k(y)\big)^\dagger \big(H_k\big)^{-1} C^{[1]}_k(A)X \\
&= \Big(\big(P^{[2]}_{n-1}(y)\big)^\dagger \big(H_{n-1}\big)^{-1} C^{[1]}_n(A) - \big(P^{[2]}_n(y)\big)^\dagger \big(H_{n-1}\big)^{-1} C^{[1]}_{n-1}(A) + I_p\Big)(A - I_p y)^{-1} X \\
&=: K^{(pc)}_{n-1}(A,y)X
\end{aligned}
$$

and $\mathcal{V}(x,y) = I_p$, so that $\mathcal{J}_{\mathcal{V}} = X$. Here we have recalled the Christoffel–Darboux formula for mixed kernels. Thus, for $n \geq 1$ we have

$$
\check{P}^{[1]}_n = \Theta_* \begin{bmatrix} C^{[1]}_{n-1}(A)X & P^{[1]}_{n-1}(x) \\ C^{[1]}_n(A)X & P^{[1]}_n(x) \end{bmatrix} = P^{[1]}_n(x) - C^{[1]}_n(A)\big(C^{[1]}_{n-1}(A)\big)^{-1} P^{[1]}_{n-1}(x),
$$

$$\check{H}_n = \Theta_* \begin{bmatrix} C^{[1]}_{n-1}(A)X & H_{n-1} \\ C^{[1]}_{n}(A)X & 0_p \end{bmatrix} = -C^{[1]}_{n}(A)\big(C^{[1]}_{n-1}(A)\big)^{-1}H_{n-1},$$

$$\big(\check{P}^{[2]}_n(x)\big)^{\dagger} = \Theta_* \begin{bmatrix} C^{[1]}_{n-1}(A)X & H_{n-1} \\ (I_p y - A)\Big(K^{(pc)}_{n-1}(A,y)+I_p\Big)X & 0_p \end{bmatrix} = \big((I_p y - A)K^{(pc)}_{n-1}(A,y) + I_p\big)\big(C^{[1]}_{n-1}(A)\big)^{-1}H_{n-1}.$$

Chapter 6

Matrix Geronimus-Uvarov transformations for matrix bi-orthogonal polynomials on the real line

In 1969 (see §1 of [143]), for the first time a massless Uvarov transformation for scalar orthogonal polynomials, finding (as Uvarov called them) general Christoffel formulas for these transformations was considered. Those results constitute a detailed version of the results presented in [144] in 1959 in Russian (see [149, 71] for more details). We now consider a transformation generated by two matrix polynomials $W_C(x), W_G(x) \in \mathbb{R}^{p\times p}[x]$, that we call Christoffel and Geronimus polynomials, respectively. This can be understood as a composition of a Geronimus transformation as discussed in the previous Chapter and a Christoffel transformation (in this order) as discussed in Chapter 4 (see also [11]).

Definition 6.1. *Given two matrix polynomials $W_C(x), W_G(x)$ of degrees N_C, N_G, respectively, such that $\sigma(W_G) \cap \operatorname{supp}(u) = \varnothing$, a matrix of linear functionals $\hat{u}$ is said to be a matrix Geronimus-Uvarov transformation of the matrix of linear functional $u = (u_{i,j}) \in (\mathcal{O}_c')^{N\times N}$, if*

$$\hat{u} W_G(x) = W_C(x) u.$$

The above definition implies that the corresponding sesquilinear forms associated to $\hat{u}$ and u satisfy

$$\left\langle P(x), Q(x) W_G^{\dagger}(x) \right\rangle_{\hat{u}} = \langle P(x) W_C(x), Q(x) \rangle_u .$$

Proposition 6.2. *The most general matrix Geronimus-Uvarov transformation is given by*

$$\hat{u} := W_C(x) u (W_G(x))^{-1} + v, \qquad v := \sum_{a=1}^{q_G} \sum_{j=1}^{s_{G,a}} \sum_{m=0}^{\kappa_{G,j}^{(a)}-1} (-1)^m \delta^{(m)}(x - x_{G,a}) W_C(x) \frac{\xi_{j,m}^{[a]}}{m!} l_{G,j}^{(a)}(x). \quad (6.1)$$

expressed in terms of the spectrum $\sigma(W_G)=\{x_{G,a}\}_{a=1}^{q_G}$, *number of Jordan blocks* $s_{G,a}$, *partial multiplicities* $\kappa^{(a)}_{G,j}$, *and corresponding adapted left root polynomials* $l^{(a)}_{G,j}(x)$ *of the matrix polynomial* $W_G(x)$ *and* $\xi^{[a]}_{j,m}\in\mathbb{C}^p$.

Observe that the matrix of functionals v is associated with the eigenvalues and left root vectors of the perturbing polynomial W_G. Also notice that we have introduced $W_C(x)$ in this term. In general we have $N_G p\geq\sum_{a=1}^{q}\sum_{i=1}^{s_a}\kappa^{(a)}_{G,j}$ and we can not ensure the equality, up to for the case of nonsingular leading coefficient.

Proposition 6.3. *The moment matrix* $\hat{M}:=\langle\chi(x),\chi(x)\rangle_{\hat{u}}$ *fulfills*

$$\hat{M}W_G(\Lambda^\top)=W_C(\Lambda)M.$$

As for the Geronimus transformation we assume that the perturbed moment matrix admits a Gauss–Borel factorization

$$\hat{M}=\hat{S}_1^{-1}\hat{H}(\hat{S}_2)^{-\dagger},$$

where $\hat{S}_1,\hat{S}_2$ are lower unitriangular block matrices and $\hat{H}$ is a diagonal block matrix

$$\hat{S}_i=\begin{bmatrix} I_p & 0_p & 0_p & \dots\\ (\hat{S}_i)_{1,0} & I_p & 0_p & \cdots\\ (\hat{S}_i)_{2,0} & (\hat{S}_i)_{2,1} & I_p & \ddots\\ & & \ddots & \ddots\end{bmatrix},\qquad i=1,2,\qquad \hat{H}=\operatorname{diag}(\hat{H}_0,\hat{H}_1,\hat{H}_2,\dots).$$

Consequently, the perturbed matrix of linear functionals gives a family of matrix bi-orthogonal polynomials

$$\hat{P}^{[i]}(x)=\hat{S}_i\chi(x),\qquad i=1,2,$$

with respect to the perturbed sesquilinear form $\langle\cdot,\cdot\rangle_{\hat{u}}$.

6.1 The resolvent and connection formulas for the matrix Geronimus-Uvarov transformation

Definition 6.4. *The resolvent matrix is given by*

$$\omega:=\hat{S}_1W_C(\Lambda)(S_1)^{-1}.$$

Proposition 6.5. i) *The resolvent matrix can also be expressed as*

$$\omega=\hat{H}\big(\hat{S}_2\big)^{-\dagger}W_G(\Lambda^\top)\big(S_2\big)^{\dagger}H^{-1}. \tag{6.2}$$

ii) *The resolvent matrix is a block band matrix with at most the first N_G block subdiagonals the main diagonal and the N_C block superdiagonals are nonzero, i.e.,*

$$\omega=\begin{bmatrix}
\omega_{0,0} & \omega_{0,1} & \cdots\;\cdots & \omega_{0,N_C-1} & I_p & 0_p & \cdots & & \\
\omega_{1,0} & \omega_{1,1} & \cdots\;\cdots & \omega_{1,N_C-1} & \omega_{1,N_C} & I_p & \ddots & & \\
\vdots & \vdots & & & \ddots & & \ddots & & \\
\omega_{N_G,0} & \omega_{N_G,1} & & & & \omega_{N_G,N_C+N_G-1} & I_p & 0_p & \\
0_p & \omega_{N_G+1,1} & & & & \omega_{N_G+1,N_G+N_C-1} & \omega_{N_G+1,N_G+N_C} & I_p & \ddots \\
\vdots & & \ddots & & & \ddots & \ddots & \ddots & \ddots
\end{bmatrix}.$$

iii) *The following connection formulas hold*

$$\hat{P}^{[1]}(x)W_C(x)=\omega P^{[1]}(x), \tag{6.3}$$

$$\big(\hat{P}^{[2]}(x)\big)^{\dagger}\hat{H}^{-1}\omega=W_G(x)\big(P^{[2]}(x)\big)^{\dagger}H^{-1}. \tag{6.4}$$

iv) *For the last subdiagonal of the resolvent we have*

$$\omega_{N_G+k,k}=\hat{H}_{N_G+k}A_{G,N_G}(H_k)^{-1}, \tag{6.5}$$

where A_{G,N_G} is the leading coefficient of the perturbing matrix polynomial $W_G(x)$.

Proof. i) Proposition 6.3 and the Gauss–Borel factorizations of the moment matrices M and $\hat{M}$ lead to

$$W_C(\Lambda)\big(S_1\big)^{-1}H\big(S_2\big)^{-\dagger}=\big(\hat{S}_1\big)^{-1}\hat{H}\big(\hat{S}_2\big)^{-\dagger}W_G(\Lambda^{\top}),$$

or, equivalently,

$$\hat{S}_1W_C(\Lambda)\big(S_1\big)^{-1}H=\hat{H}\big(\hat{S}_2\big)^{-\dagger}W_G(\Lambda^{\top})\big(S_2\big)^{\dagger}$$

and the result follows.

ii) From its definition, the resolvent matrix is a lower generalized Hessenberg block matrix, with N_C nonzero superdiagonals. However, from (6.2) we deduce that it is an upper generalized block Hessenberg matrix with N_G nonzero subdiagonals. As a conclusion, we get the band structure.

iii) From the definition of the resolvent we deduce (6.3). From (6.2) we get (6.4).

iv) It is a direct observation from (6.2).

■

Connection formulas (6.3) and (6.4) can be written as

$$\hat{P}_n^{[1]}(x)W_C(x)=P_{n+N_C}^{[1]}(x)+\sum_{k=n-N_G}^{n+N_C-1}\omega_{n,k}P_k^{[1]}(x),$$

$$W_G(x)\big(P_n^{[2]}(x)\big)^{\dagger}(H_n)^{-1}=\sum_{k=n-N_C}^{n+N_G}\big(\hat{P}_k^{[2]}(x)\big)^{\dagger}(\hat{H}_k)^{-1}\omega_{k,n}. \tag{6.6}$$

Proposition 6.6. *The matrix Geronimus-Uvarov transformation of the second kind functions reads as*

$$\hat{C}^{[1]}(z)W_G(z)-\begin{bmatrix}\left(\hat{H}\left(\hat{S}_2\right)^{-\dagger}\right)_{[N_G]}\mathcal{B}_G(\chi(z))_{[N_G]}\\ 0_p\\ \vdots\end{bmatrix}=\omega C^{[1]}(z), \tag{6.7}$$

$$\left(\hat{C}^{[2]}(z)\right)^{\dagger}\hat{H}^{-1}\omega=W_C(z)\left(C^{[2]}(z)\right)^{\dagger}H^{-1}-\left[(\chi(z))^{\dagger}_{[N_C]}\mathcal{B}_C\left(S_1\right)^{-1}_{[N_C]},0_p,\dots\right]. \tag{6.8}$$

Proof. From (1.9) we can write

$$\begin{aligned}\omega C^{[1]}(z)-\hat{C}^{[1]}(z)W_G(z)&=\omega\left\langle P^{[1]}(x),\frac{I_p}{z-x}\right\rangle_u-\left\langle \hat{P}^{[1]}(x),\frac{I_p}{z-x}\right\rangle_{\hat{u}}W_G(z)\\ &=\left\langle \hat{P}^{[1]}(x)W_C(x),\frac{I_p}{z-x}\right\rangle_u-\left\langle \hat{P}^{[1]}(x),\frac{I_p}{z-x}\right\rangle_{\hat{u}}W_G(z)\\ &=\left\langle \hat{P}^{[1]}(x),\frac{W_G(x)^{\dagger}}{z-x}\right\rangle_u-\left\langle \hat{P}^{[1]}(x),\frac{I_p}{z-x}\right\rangle_{\hat{u}}W_G(z)\\ &=-\left\langle \hat{P}^{[1]}(x),\frac{(W_G(x)-W_G(z))^{\dagger}}{x-z}\right\rangle_{\hat{u}}=-\hat{S}_1\left\langle \chi(x),\frac{(W_G(x)-W_G(z))^{\dagger}}{x-z}\right\rangle_{\hat{u}}.\end{aligned}$$

Using the fact that $\frac{(W_G(x)-W_G(z))^{\dagger}}{x-z}$ is a matrix polynomial of degree N_G-1 together with (5.10), we get (6.7). For (6.8) we observe that from (6.6)

$$\begin{aligned}\sum_{k=n-N_C}^{n+N_G}\hat{C}_k^{[2]\dagger}(z)(\hat{H}_k)^{-1}\omega_{k,n}&=\left\langle\frac{I_p}{z-x},P_n^{[2]}(x)W_G(x)^{\dagger}\right\rangle_{\hat{u}}H_n^{-1}\\ &=\left\langle\frac{W_C(x)-W_C(z)}{z-x},P_n^{[2]}(x)\right\rangle_u H_n^{-1}+\left\langle\frac{W_C(z)}{z-x},P_n^{[2]}(x)\right\rangle_u H_n^{-1}\\ &=-\left\langle\frac{W_C(z)-W_C(x)}{z-x},P_n^{[2]}(x)\right\rangle_u H_n^{-1}+W_C(z)(C_n^{[2]})^{\dagger}(z)H_n^{-1},\end{aligned}$$

and since $\frac{W_C(z)-W_C(x)}{z-x}$ is a polynomial of degree N_C-1, then using (5.9)

$$\left\langle\frac{W_C(z)-W_C(x)}{z-x},P_n^{[2]}(x)\right\rangle_u=\begin{cases}0_p, & \text{if } n>N_C-1,\\ [\chi(z)^{\dagger}]_{[N_C]}\mathcal{B}_C\left\langle[\chi(x)]_{[N_C]},P_n^{[2]}(x)\right\rangle_u, & \text{if } n\leq N_C-1.\end{cases}$$

Thus, we get the result. ∎

Remark 6.7. *We can understand a matrix Geronimus-Uvarov transformation as a composition of a Geronimus transformation (the first step) and next a Christoffel transformation, in terms of the corresponding matrices of linear functionals $u\mapsto\check{u}\mapsto\hat{u}$, where*

$$\check{u}W_G=u,\qquad \hat{u}=W_C\check{u}.$$

At the level of the moment matrices we get

$$\check{M}W_G(\Lambda^\top) = M, \qquad \hat{M} = W_C(\Lambda)\check{M}.$$

Here, we will assume that the linear spectral transformation can be performed in two steps, i.e. that $\check{M}$ has a Gauss–Borel factorization[1]*. Thus we can write*

$$\big(\check{S}_1\big)^{-1}\check{H}\big(\check{S}_2\big)^{-\dagger}W_G(\Lambda^\top) = \big(S_1\big)^{-1}H\big(S_2\big)^{-\dagger}, \quad \big(\hat{S}_1\big)^{-1}\hat{H}\big(\hat{S}_2\big)^{-\dagger} = W_C(\Lambda)\big(\check{S}_1\big)^{-1}\check{H}\big(\check{S}_2\big)^{-\dagger}.$$

Therefore, we have the partial resolvent

$$\begin{aligned}\omega_G &:= \check{S}_1\big(S_1\big)^{-1} = \check{H}\big(\check{S}_2\big)^{-\dagger}W_G(\Lambda^\top)S_2H^{-1},\\ \omega_C &:= \hat{S}_1W_C(\Lambda)\big(\check{S}_1\big)^{-1} = \hat{H}\big(\hat{S}_2\big)^{-\dagger}\big(\check{S}_2\big)^{\dagger}\check{H}^{-1}.\end{aligned}$$

We see that ω_G is a lower unitriangular block semi-infinite matrix with only the first N_G subdiagonals non-zero, and ω_C is an upper triangular block semi-infinite matrix with only the N_C first superdiagonals non-zero. The resolvent ω results from the composition of both transformations, so that $\omega = \omega_C\omega_G$. We also notice that ω_G, being unitriangular, is nonsingular, and has nonsingular coefficients as its last nonzero subdiagonal. Regarding ω_C, the block entries in its highest nonzero superdiagonal are identity matrices I_p, and it has on the main diagonal nonsingular matrices as entries and, consequently, it has an inverse. Thus, the equation $\omega X = 0$ implies $X = 0$.

6.1.1 Matrix Geronimus-Uvarov transformation and Christoffel–Darboux kernels

Definition 6.8. *We introduce the resolvent wing matrices*

$$\Omega^G[n] := \begin{cases} \begin{bmatrix} \omega_{n,n-N_G} & \cdots & \cdots & \omega_{n,n-1} \\ 0_p & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0_p & \cdots & 0_p & \omega_{n+N_G-1,n-1} \end{bmatrix} \in \mathbb{C}^{N_Gp\times N_Gp}, & n \geq N_G, \\[2ex] \begin{bmatrix} \omega_{n,0} & \cdots & \cdots & \omega_{n,n-1} \\ \vdots & & & \vdots \\ \omega_{N_G,0} & & & \omega_{N_G,n-1} \\ 0_p & \ddots & & \\ \vdots & \ddots & \ddots & \vdots \\ 0_p & \cdots & 0_p & \omega_{n+N_G-1,n-1} \end{bmatrix} \in \mathbb{C}^{N_Gp\times np}, & n < N_G, \end{cases}$$

[1] We already assume that M and $\hat{M}$ do have such a factorization.

$$\Omega^C[n] := \begin{cases} \begin{bmatrix} I_p & 0_p & \ddots & 0_p & 0_p \\ \omega_{n-N_C+1,n} & I_p & \ddots & 0_p & 0_p \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & I_p & 0_p \\ \omega_{n-1,n} & \dots & & \omega_{n-1,n+N_C-2} & I_p \end{bmatrix} \in \mathbb{C}^{N_C p\times N_C p}, & n \geq N_C, \\ \begin{bmatrix} \omega_{0,n} & \dots & \omega_{0,N_C-1} & I_p & \dots & 0 \\ \vdots & & & & \ddots & \vdots \\ \omega_{n-1,n} & \dots & & \dots & \omega_{n-1,n+N_C-2} & I_p \end{bmatrix} \in \mathbb{C}^{np\times N_C p}, & n < N_C. \end{cases}$$

Theorem 6.9 (Connection formulas for the Christoffel–Darboux kernels). *For $m_G = \min(n,N_G)$ and $m_C = \min(n,N_C)$, the perturbed and original Christoffel–Darboux kernels are connected as follows*

$$\begin{aligned} &\hat{K}_{n-1}(x,y)W_C(x) = W_G(y)K_{n-1}(x,y) \\ &- \begin{bmatrix} \left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} & \cdots & \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1} \end{bmatrix} \begin{bmatrix} 0_{m_C p\times m_G p} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_G p\times N_C p} \end{bmatrix} \begin{bmatrix} P^{[1]}_{n-m_G}(x) \\ \vdots \\ P^{[1]}_{n+N_C-1}(x) \end{bmatrix}. \end{aligned} \tag{6.9}$$

For $n \geq N_G$, the mixed Christoffel–Darboux kernels satisfy

$$\begin{aligned} &\hat{K}^{(pc)}_{n-1}(x,y)W_G(x) = W_G(y)K^{(pc)}_{n-1}(x,y) + \mathcal{V}_G(x,y) \\ &- \begin{bmatrix} \left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} & \cdots & \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1} \end{bmatrix} \begin{bmatrix} 0_{m_C p\times N_G p} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_G p\times N_C p} \end{bmatrix} \begin{bmatrix} C^{[1]}_{n-N_G}(x) \\ \vdots \\ C^{[1]}_{n+N_C-1}(x) \end{bmatrix}. \end{aligned} \tag{6.10}$$

Where $\mathcal{V}_G(x,y)$ is given in Definition 1.42.

Proof. For the first connection formulas (6.9) we consider

$$\mathcal{K}_{n-1}(x,y) := \begin{bmatrix} (\hat{P}^{[2]}_0(y))^{\dagger}(\hat{H}_0)^{-1} & \cdots & (\hat{P}^{[2]}_{n-1}(y))^{\dagger}(\hat{H}_{n-1})^{-1} \end{bmatrix} \omega_{[n]} \begin{bmatrix} P^{[1]}_0(x) \\ \vdots \\ P^{[1]}_{n-1}(x) \end{bmatrix},$$

and we can compute it in two different ways. From (6.3) we get that

$$\omega_{[n]}\begin{bmatrix} P_0^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x)\end{bmatrix} = \begin{bmatrix} \hat{P}_0^{[1]}(x) \\ \vdots \\ \hat{P}_{n-1}^{[1]}(x)\end{bmatrix} W_C(x) - \tilde{\Omega}^C[n]\begin{bmatrix} P_n^{[1]}(x) \\ \vdots \\ P_{n+N_C-1}^{[1]}(x)\end{bmatrix},$$

where we have used the notation

$$\tilde{\Omega}^C[n] := \begin{cases} \begin{bmatrix} 0_{(n-N_c)p\times N_c p} \\ \Omega^C[n]\end{bmatrix}, & n \geq N_C, \\ \Omega^C[n], & n < N_C. \end{cases}$$

Therefore, for $m_C = \min(N_C, n)$, we can write

$$\mathcal{K}_{n-1}(x,y) = \hat{K}_{n-1}(x,y)W_C(x) - \Big[\big(\hat{P}^{[2]}_{n-m_C}(y)\big)^\dagger (\hat{H}_{n-m_C})^{-1}, \dots, \big(\hat{P}^{[2]}_{n-1}(y)\big)^\dagger (\hat{H}_{n-1})^{-1}\Big]\Omega^C[n]\begin{bmatrix} P_n^{[1]}(x) \\ \vdots \\ P_{n+N_C-1}^{[1]}(x)\end{bmatrix}.$$

For $m_G = \min(n, N_G)$, (6.4) leads to

$$\mathcal{K}_{n-1}(x,y) = W_G(y)K_{n-1}(x,y) - \Big[\big(\hat{P}^{[2]}_{n}(y)\big)^\dagger (\hat{H}_{n})^{-1}, \dots, \big(\hat{P}^{[2]}_{n+N_G-1}(y)\big)^\dagger (\hat{H}_{n+N_G-1})^{-1}\Big]\Omega^G[n]\begin{bmatrix} P_{n-m_G}^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x)\end{bmatrix}.$$

Hence, we obtain

$$\begin{aligned} &W_G(y)K_{n-1}(x,y) - \Big[\big(\hat{P}^{[2]}_{n}(y)\big)^\dagger (\hat{H}_{n})^{-1}, \dots, \big(\hat{P}^{[2]}_{n+N_G-1}(y)\big)^\dagger (\hat{H}_{n+N_G-1})^{-1}\Big]\Omega^G[n]\begin{bmatrix} P_{n-m_G}^{[1]}(x) \\ \vdots \\ P_{n-1}^{[1]}(x)\end{bmatrix} \\ &= \hat{K}_{n-1}(x,y)W_C(x) - \Big[\big(\hat{P}^{[2]}_{n-m_C}(y)\big)^\dagger (\hat{H}_{n-m_C})^{-1}, \dots, \big(\hat{P}^{[2]}_{n-1}(y)\big)^\dagger (\hat{H}_{n-1})^{-1}\Big]\Omega^C[n]\begin{bmatrix} P_n^{[1]}(x) \\ \vdots \\ P_{n+N_C-1}^{[1]}(x)\end{bmatrix}, \end{aligned}$$

and (6.9) follows. Next, we consider

$$\mathcal{K}^{(pc)}_{n-1}(x,y) := \Big[\big(\hat{P}^{[2]}_{0}(y)\big)^\dagger (\hat{H}_{0})^{-1}, \dots, \big(\hat{P}^{[2]}_{n-1}(y)\big)^\dagger (\hat{H}_{n-1})^{-1}\Big]\omega_{[n]}\begin{bmatrix} C_0^{[1]}(x) \\ \vdots \\ C_{n-1}^{[1]}(x)\end{bmatrix},$$

which can be computed in two different ways. First, using (6.7) we get

$$
\begin{aligned}
\mathcal{K}^{(pc)}_{n-1}(x,y) &= \left[\left(\hat{P}^{[2]}_{0}(y)\right)^{\dagger}(\hat{H}_0)^{-1},\dots,\left(\hat{P}^{[2]}_{n-1}(y)\right)^{\dagger}(\hat{H}_{n-1})^{-1}\right] \\
&\quad\times\left(\begin{bmatrix}\hat{C}^{[1]}_{0}(x)W_G(x)\\ \vdots \\ \hat{C}^{[1]}_{n-1}(x)W_G(x)\end{bmatrix} - \left(\hat{H}\left(\hat{S}_2\right)^{-\dagger}\right)_{[n,N_G]}\mathcal{B}_G(\chi(z))_{[N_G]} - \tilde{\Omega}^{C}[n]\begin{bmatrix}C^{[1]}_{n}(x)\\ \vdots \\ C^{[1]}_{n+N_C-1}(x)\end{bmatrix}\right) \\
&= \hat{K}^{(pc)}_{n-1}(x,y)W_G(x) - \left((\chi(y))_{[n]}\right)^{\dagger}\left(\left(\hat{S}_2\right)^{\dagger}\hat{H}^{-1}\right)_{[n]}\left(\hat{H}\left(\hat{S}_2\right)^{-\dagger}\right)_{[n,N_G]}\mathcal{B}_G(\chi(x))_{[N_G]} \\
&\qquad - \left[\left(\hat{P}^{[2]}_{0}(y)\right)^{\dagger}(\hat{H}_0)^{-1},\dots,\left(\hat{P}^{[2]}_{n-1}(y)\right)^{\dagger}(\hat{H}_{n-1})^{-1}\right]\tilde{\Omega}^{C}[n]\begin{bmatrix}C^{[1]}_{n}(x)\\ \vdots \\ C^{[1]}_{n+N_C-1}(x)\end{bmatrix},
\end{aligned}
$$

where again $\left(\hat{H}\left(\hat{S}_2\right)^{-\dagger}\right)_{[n,N_G]}$ denotes the truncation to the n first block rows and first N_G block columns of $\hat{H}\left(\hat{S}_2\right)^{-\dagger}$. For $n \geq N_G$ this simplifies to

$$
\begin{aligned}
\mathcal{K}^{(pc)}_{n-1}(x,y) &= \hat{K}^{(pc)}_{n-1}(x,y)W_G(x) - \left((\chi(y))_{[N_G]}\right)^{\dagger}\mathcal{B}_G(\chi(x))_{[N_G]} \\
&\qquad - \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1},\dots,\left(\hat{P}^{[2]}_{n-1}(y)\right)^{\dagger}(\hat{H}_{n-1})^{-1}\right]\Omega^{C}[n]\begin{bmatrix}C^{[1]}_{n}(x)\\ \vdots \\ C^{[1]}_{n+N_C-1}(x)\end{bmatrix}.
\end{aligned}
$$

Second, if we use (6.4), it leads to

$$
\mathcal{K}^{(pc)}_{n-1}(x,y) = W_G(y)K^{(pc)}_{n-1}(x,y) - \left[\left(\hat{P}^{[2]}_{n}(y)\right)^{\dagger}(\hat{H}_n)^{-1},\dots,\left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\Omega^{G}[n]\begin{bmatrix}C^{[1]}_{n-m_G}(x)\\ \vdots \\ C^{[1]}_{n-1}(x)\end{bmatrix},
$$

and, consequently, we obtain

$$
\begin{aligned}
&\hat{K}^{(pc)}_{n-1}(x,y)W_G(x) - \mathcal{V}_G(x,y) - \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1},\dots,\left(\hat{P}^{[2]}_{n-1}(y)\right)^{\dagger}(\hat{H}_{n-1})^{-1}\right]\Omega^{C}[n]\begin{bmatrix}C^{[1]}_{n}(x)\\ \vdots \\ C^{[1]}_{n+N_C-1}(x)\end{bmatrix} \\
&\qquad = W_G(y)K^{(pc)}_{n-1}(x,y) - \left[\left(\hat{P}^{[2]}_{n}(y)\right)^{\dagger}(\hat{H}_n)^{-1},\dots,\left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\Omega^{G}[n]\begin{bmatrix}C^{[1]}_{n-N_G}(x)\\ \vdots \\ C^{[1]}_{n-1}(x)\end{bmatrix}.
\end{aligned}
$$

■

6.2 Spectral properties of the first family of perturbed second kind functions

In the next sections we will assume that the perturbing polynomials are monic, i.e. $W_G(x) = I_p x^N + \sum_{k=0}^{N-1} A_{G,k} x^k$, $W_C(x) = I_p x^N + \sum_{k=0}^{N-1} A_{C,k} x^k$. The corresponding spectral data, eigenvalues, algebraic multiplicities, partial multiplicities, left and right root polynomials, and Jordan pairs and triples will have a subindex C or G to indicate to which polynomial $W_C(x)$ or $W_G(x)$ they are linked to.

Definition 6.10. *Let $x_{G,a}$ be an eigenvalue of $W_G(x)$. For $z \neq x_{G,a}$, we introduce the $p \times p$ matrices*

$$\hat{C}^{(a)}_{n;i}(z) := \mathtt{J}^{(i)}_{\hat{P}^{[1]}_n W_C}(x_{G,a}) \mathcal{X}^{(a)}_i \mathcal{L}^{(a)}_{G,i} \begin{bmatrix} \frac{I_p}{(z-x_{G,a})^{\kappa^{(a)}_i}} \\ \vdots \\ \frac{I_p}{z-x_{G,a}} \end{bmatrix}, \tag{6.11}$$

where $i = 1, \dots, s_{G,a}$, and the matrices $\mathcal{X}^{(a)}_i$, $\mathcal{L}^{(a)}_{G,i}$ are defined as in (5.14) and (5.15).

Proposition 6.11. *For $z \notin \operatorname{supp}(u) \cup \sigma(W_G)$, the following expression*

$$\hat{C}^{[1]}_n(z) = \left\langle \hat{P}^{[1]}_n(x), \frac{I_p}{z-x} \right\rangle_{W_C u W_G^{-1}} + \sum_{a=1}^{q_G} \sum_{i=1}^{s_{G,a}} \hat{C}^{(a)}_{n;i}(z)$$

holds.

Proof. From Definition 1.9

$$\begin{aligned} \hat{C}^{[1]}_n(z) &= \left\langle \hat{P}^{[1]}_n(x), \frac{I_p}{z-x} \right\rangle_{\hat{u}} \\ &= \left\langle \hat{P}^{[1]}_n(x), \frac{I_p}{z-x} \right\rangle_{W_C u W_G^{-1}} + \sum_{a=1}^{q_G} \sum_{i=1}^{s_{G,a}} \sum_{m=0}^{\kappa^{(a)}_{G,i}-1} \left(\hat{P}^{[1]}_n(x) W_C(x) \frac{\xi^{[a]}_{i,m}}{m!} \frac{l^{(a)}_{G,i}(x)}{(z-x)} \right)^{(m)}_{x=x_{G,a}}. \end{aligned}$$

Now, taking into account that

$$\left(\hat{P}^{[1]}_n(x) W_C(x) \frac{\xi^{[a]}_{i,m}}{m!} \frac{l^{(a)}_{G,i}(x)}{(z-x)} \right)^{(m)}_{x=x_{G,a}} = \sum_{k=0}^{m} \left(\hat{P}^{[1]}_n(x) W_C(x) \frac{\xi^{[a]}_{i,m} l^{(a)}_{G,i}(x)}{(m-k)!} \right)^{(m-k)}_{x=x_{G,a}} \frac{1}{(z-x_{G,a})^{k+1}},$$

we deduce the result. ■

Lemma 6.12. *The function $\hat{C}^{(a)}_{n;i}(x) W_G(x) r^{(b)}_{G,j}(x) \in \mathbb{C}^p[x]$ satisfies*

$$\hat{C}^{(a)}_{n;i}(x) W_G(x) r^{(b)}_{G,j}(x) = \tag{6.12}$$

$$\begin{cases}\mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i\begin{bmatrix}(x-x_{G,a})^{\kappa^{(a)}_{G,\max(i,j)}-\kappa^{(a)}_{G,i}}\\ \vdots\\ (x-x_{G,a})^{\kappa^{(a)}_{G,\max(i,j)}-1}\end{bmatrix}w^{(a)}_{G,i,j}(x)+(x-x_{G,a})^{\kappa^{(a)}_{G,j}}T^{(a,a)}(x), & a=b,\\ (x-x_{G,b})^{\kappa^{(b)}_{G,j}}T^{(a,b)}(x), & a\neq b,\end{cases}\tag{6.13}$$

where the $\mathbb{C}^p$-valued function $T^{(a,b)}(x)$ is analytic at $x=x_{G,b}$ and, in particular, $T^{(a,a)}(x)\in\mathbb{C}^p[x]$.

Proof. The function $\hat{C}^{(a)}_{n;i}(x)W_G(x)r^{(b)}_{G,j}(x)\in\mathbb{C}^p[x]$, with $a\neq b$, as (1.5) informs us about, is such that

$$\begin{aligned}\hat{C}^{(a)}_{n;i}(x)W_G(x)r^{(b)}_{G,j}(x)&=\mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i\mathcal{L}^{(a)}_{G,i}\begin{bmatrix}\dfrac{I_p}{(x-x_{G,a})^{\kappa^{(a)}_{G,i}}}\\ \vdots\\ \dfrac{I_p}{x-x_{G,a}}\end{bmatrix}W_G(x)r^{(b)}_{G,j}(x)\\ &=(x-x_{G,b})^{\kappa^{(b)}_{G,j}}T^{(a,b)}(x),\end{aligned}$$

where $T^{(a,b)}(x)$ is an $\mathbb{C}^p$-valued analytic function at $x=x_{G,b}$. (6.11) and Lemma (5.12) yield

$$\hat{C}^{(a)}_{n;i}(x)W_G(x)r^{(a)}_{G,j}(x)=\mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i\mathcal{L}^{(a)}_{G,i}\begin{bmatrix}\dfrac{I_p}{(x-x_{G,a})^{\kappa^{(a)}_{G,i}}}\\ \vdots\\ \dfrac{I_p}{x-x_{G,a}}\end{bmatrix}W_G(x)r^{(a)}_{G,j}(x)$$

$$=\mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i\begin{bmatrix}\dfrac{1}{(x-x_{G,a})^{\kappa^{(a)}_{G,i}}}\\ \vdots\\ \dfrac{1}{x-x_{G,a}}\end{bmatrix}l^{(a)}_{G,i}(x)W_G(x)r^{(a)}_{G,j}(x)+(x-x_{G,a})^{\kappa^{(a)}_{G,j}}\mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i T^{(a,a)}(x),$$

for some $T^{(a,a)}(x)\in\mathbb{C}^p[x]$. Hence, using Proposition 1.34 we get

$$\begin{aligned}\hat{C}^{(a)}_{n;i}(x)W_G(x)r^{(a)}_{G,j}(x)&=\mathtt{J}^{(i)}_{G,\check{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i\begin{bmatrix}(x-x_{G,a})^{\kappa^{(a)}_{G,\max(i,j)}-\kappa^{(a)}_{G,i}}\\ \vdots\\ (x-x_{G,a})^{\kappa^{(a)}_{G,\max(i,j)}-1}\end{bmatrix}\\ &\quad\times\Big(w^{(a)}_{G,i,j;0}+w^{(a)}_{G,i,j;1}(x-x_{G,a})+\cdots+w^{(a)}_{G,i,j;\kappa^{(a)}_{G,\min(i,j)}+N_G-2}(x-x_{G,a})^{\kappa^{(a)}_{G,\min(i,j)}+N_G-2}\Big).\\ &\quad+(x-x_{G,a})^{\kappa^{(a)}_{G,j}}\mathtt{J}^{(i)}_{G,\check{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i T^{(a,a)}(x),\end{aligned}$$

and the result follows. ∎

Lemma 6.13. *The following relations hold*

$$\left(\hat{C}_n^{[1]}(z)W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)} = \sum_{i=1}^{s_{G,a}} \left(\hat{C}_{n;i}^{(a)}(z)W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)}, \tag{6.14}$$

for $m = 0,\dots,\kappa_{G,j}^{(a)}-1$.

Proof. For $z \notin \operatorname{supp}(u)\cup\sigma(W_G)$, Proposition 6.11 leads to

$$\begin{aligned}&\left(\hat{C}_n^{[1]}(z)W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)} = \\ &\qquad \left(\left\langle \hat{P}_n^{[1]}(x), \frac{I_p}{z-x}\right\rangle_{W_C u W_G^{-1}} W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)} + \sum_{b=1}^{q_G}\sum_{i=1}^{s_{G,b}} \left(\hat{C}_{n;i}^{(b)}(z)W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)}.\end{aligned}$$

Taking into account $\sigma(W)\cap\operatorname{supp}(u)=\varnothing$, the derivatives of the Cauchy kernel $1/(z-x)$ are analytic functions at $z = x_{G,a}$. Hence,

$$\begin{aligned}\left(\left\langle \hat{P}_n^{[1]}(x), \frac{I_p}{z-x}\right\rangle_{W_C u W_G^{-1}} W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)} &= \left\langle \hat{P}_n^{[1]}(x), \left(\frac{[W_G(z)r_{G,j}^{(a)}(z)]^\dagger}{z-x}\right)_{z=x_{G,a}}^{(m)}\right\rangle_{W_C u W_G^{-1}} \\ &= \left\langle \hat{P}_n^{[1]}(x), \left[\sum_{k=0}^{m}\binom{m}{k}\left(W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(k)} \frac{(-1)^{m-k}(m-k)!}{(x_{G,a}-x)^{m-k+1}}\right]^\dagger\right\rangle_{W_C u W_G^{-1}} \\ &= 0_{p\times 1},\end{aligned}$$

for $m = 0,\dots,\kappa_{G,j}^{(a)}-1$. From (6.12) we know that if $b\neq a$, then $\hat{C}_{n;i}^{(b)}(z)W_G(z)r_{G,j}^{(a)}(z)$ has a zero at $z = x_{G,a}$ of order $\kappa_{G,j}^{(a)}$, i.e.

$$\left(\hat{C}_{n;i}^{(b)}(z)W_G(z)r_{G,j}^{(a)}(z)\right)_{z=x_{G,a}}^{(m)} = 0, \qquad b\neq a,$$

for $m = 0,\dots,\kappa_{G,j}^{(a)}-1$. ∎

Let $\mathcal{W}_{G,i,j}^{(a)}$, $\mathcal{W}_{G,i}^{(a)}$, $\mathcal{T}_G^{(a)}$, and $\mathcal{T}^{(a)}$ be as in Definition 5.15 but now corresponding to the zeros of $W_G(x)$. Then,

Proposition 6.14. *The following formulas*

$$\mathcal{J}^{(j)}_{G,\hat{C}_n^{[1]}W_G}(x_{G,a}) = \sum_{i=1}^{s_{G,a}} \mathcal{J}^{(j)}_{G\hat{C}_{n;i}W_G}(x_{G,a}), \qquad \mathcal{J}_{G,\hat{C}_n^{[1]}W_G}(x_{G,a}) = \sum_{i=1}^{s_{G,a}} \mathcal{J}_{G,\hat{C}_{n;i}W_G}(x_{G,a}), \tag{6.15}$$

$$\mathcal{J}^{(j)}_{G,\hat{C}_{n;i}W_G}(x_{G,a}) = \mathbf{J}^{(i)}_{G,\hat{P}_n^{[1]}W_C}(x_{G,a})\mathcal{X}_i^{(a)}\mathcal{W}_{G,i,j}^{(a)}, \quad \mathcal{J}_{G,\hat{C}_{n;i}W_G}(x_{G,a}) = \mathbf{J}^{(i)}_{G,\hat{P}_n^{[1]}W_C}(x_{G,a})\mathcal{X}_i^{(a)}\mathcal{W}_{G,i}^{(a)}, \tag{6.16}$$

$$\mathcal{J}_{G,\hat{C}_n^{[1]}W_G}(x_{G,a}) = \mathbf{J}_{G,\hat{P}_n^{[1]}W_C}(x_{G,a})\mathcal{T}_G^{(a)}, \qquad \mathcal{J}_{G,\hat{C}_n^{[1]}W_G} = \mathbf{J}_{G,\hat{P}_n^{[1]}W_C}\mathcal{T}_G, \tag{6.17}$$

hold.

Proof. (6.15) is deduced from (6.14). To show (6.16), notice that according to (6.12), for $m = 0,\dots,\kappa^{(a)}_{G,j}-1$,

$$\left(\hat{C}^{(a)}_{n;i}(x)W_G(x)r^{(a)}_{G,j}(x)\right)^{(m)}_{x=x_{G,a}} = \mathtt{J}^{(i)}_{\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i \begin{bmatrix} \left((x-x_{G,a})^{\kappa^{(a)}_{G,\max(i,j)}-\kappa^{(a)}_{G,i}} w^{(a)}_{G,i,j}(x)\right)^{(m)}_{x=x_{G,a}} \\ \vdots \\ \left((x-x_{G,a})^{\kappa^{(a)}_{G,\max(i,j)}-1} w^{(a)}_{G,i,j}(x)\right)^{(m)}_{x=x_{G,a}} \end{bmatrix}.$$

For (6.17), just notice that from (6.16) and (6.15)

$$\mathcal{J}^{(j)}_{G,\hat{C}^{[1]}_n W_G}(x_{G,a}) = \sum_{i=1}^{s_{G,a}} \mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{X}^{(a)}_i \mathcal{W}^{(a)}_{G,i,j}, \quad \mathcal{J}_{G,\hat{C}^{[1]}_n W_G}(x_{G,a}) = \sum_{i=1}^{s_{G,a}} \mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n}(x_{G,a})\mathcal{X}^{(a)}_i \mathcal{W}^{(a)}_{G,i}.$$

Recalling (5.19) we have

$$\begin{aligned} \mathcal{J}_{G,\hat{C}^{[1]}_n W_G}(x_{G,a}) &= \sum_{i=1}^{s_{G,a}} \mathtt{J}^{(i)}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{T}^{(a)}_{G,i} \\ &= \mathtt{J}_{G,\hat{P}^{[1]}_n W_C}(x_{G,a})\mathcal{T}^{(a)}_G, \end{aligned}$$

and, in a similar way, we get the other relation. ■

6.3 Spectral Christoffel–Geronimus–Uvarov formulas

We assume that the leading terms of both perturbing polynomials, A_{G,N_G} and A_{C,N_C}, are nonsingular.

Discussion for $n \geq N_G$

Remark 6.15. *In Corollary 6.36, taking into account mixed spectral/nonspectral approach, we will check that if $A_{G,N}$ is nonsingular, then we have*

$$\begin{vmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathtt{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathtt{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G \end{vmatrix} \neq 0.$$

Proposition 6.16. *For $n \geq N_G$, the resolvent can be expressed as follows*

$$\left[\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\right] = -\left[\mathcal{J}_{C,P^{[1]}_{n+N_C}}, \mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathtt{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G\right] \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathtt{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathtt{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G \end{bmatrix}^{-1}.$$

Proof. The connection formula (6.7) gives for $n \geq N_G$

$$\hat{C}_n^{[1]}(x)W_G(x) = \sum_{k=n-N_G}^{n+N_C} \omega_{n,k} C_k^{[1]}(x).$$

We deduce that

$$\mathcal{J}_{G,\hat{C}_n^{[1]}W_G} = \begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{G,C^{[1]}_{n-N_G}}\\ \vdots \\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}}\end{bmatrix} + \mathcal{J}_{G,C^{[1]}_{n+N_C}}.$$

Now, from (6.3),

$$\mathbf{J}_{G,\hat{P}_n^{[1]}W_C} = \begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathbf{J}_{G,P^{[1]}_{n-N_G}}\\ \vdots \\ \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\end{bmatrix} + \mathbf{J}_{G,P^{[1]}_{n+N_C}}.$$

Recalling (6.17) we obtain

$$\begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{G,C^{[1]}_{n-N_C}}\\ \vdots \\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}}\end{bmatrix} + \mathcal{J}_{G,C^{[1]}_{n+N_C}} = \begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G\\ \vdots \\ \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix} + \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G.$$

It reads

$$\begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G\\ \vdots \\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix}_{(N_G+N_C)p\times N_G p} = -\big(\mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G\big).$$

From (6.3) we also get

$$\begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n-N_G}}\\ \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}}\end{bmatrix}_{(N_G+N_C)p\times N_C p} + \mathcal{J}_{C,P^{[1]}_{n+N_C}} = 0. \tag{6.18}$$

Therefore,

$$\begin{bmatrix}\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G\\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix} = -\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n+N_C}}, \mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G\end{bmatrix}.$$

■

Theorem 6.17 (Spectral Christoffel–Geronimus–Uvarov formulas)**.** *For monic perturbations, when $n \geq N_G$, we have the following last quasi-determinant expressions for the perturbed bi-orthogonal matrix polynomials and its matrix norms*

$$\hat{P}_n^{[1]}(x)W_C(x) = \Theta_* \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G & P^{[1]}_{n-N_G}(x) \\ \vdots & \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C}} & \mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G & P^{[1]}_{n+N_C}(x) \end{bmatrix},$$

$$\hat{H}_n = \Theta_* \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G & H_{n-N_G} \\ \mathcal{J}_{C,P^{[1]}_{n-N_G+1}} & \mathcal{J}_{G,C^{[1]}_{n-N_G+1}} - \mathbf{J}_{G,P^{[1]}_{n-N_G+1}}\mathcal{T}_G & 0_p \\ \vdots & \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C}} & \mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G & 0_p \end{bmatrix},$$

$$\left(\hat{P}_n^{[2]}(y)\right)^\dagger = -\Theta_* \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G & H_{n-N_G} \\ \mathcal{J}_{C,P^{[1]}_{n-N_G+1}} & \mathcal{J}_{G,C^{[1]}_{n-N_G+1}} - \mathbf{J}_{G,P^{[1]}_{n-N_G+1}}\mathcal{T}_G & 0_p \\ \vdots & \vdots & \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G & 0_p \\ W_G(y)\mathcal{J}_{C,K_{n-1}}(y) & W_G(y)\left(\mathcal{J}_{G,K^{(pc)}_{n-1}}(y) - \mathbf{J}_{G,K_{n-1}}(y)\mathcal{T}_G\right) + \mathcal{J}_{G,\mathcal{V}}(y) & 0_p \end{bmatrix}.$$

Proof. From (6.3)

$$\hat{P}_n^{[1]}(x)W_C(x) = P^{[1]}_{n+N_C}(x) + \left[\omega_{n,n-N_G}, \dots, \omega_{n,n+N_C-1}\right] \begin{bmatrix} P^{[1]}_{n-N_G}(x) \\ \vdots \\ P^{[1]}_{n+N_C-1}(x) \end{bmatrix},$$

and applying Proposition 6.16 we get

$$\begin{aligned} \hat{P}_n^{[1]}(x)W_C(x) &= P^{[1]}_{n+N_C}(x) \\ &- \left[\mathcal{J}_{C,P^{[1]}_{n+N_C}}, \mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G\right] \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G \end{bmatrix}^{-1} \begin{bmatrix} P^{[1]}_{n-N_G}(x) \\ \vdots \\ P^{[1]}_{n+N_C-1}(x) \end{bmatrix}, \end{aligned}$$

and the result is proven. From (6.5) we deduce

$$\hat{H}_n = \omega_{n,n-N_G} H_{n-N_G}.$$

According to Proposition 6.16,

$$\omega_{n,n-N_G} = -\left[\mathcal{J}_{C,P^{[1]}_{n+N_C}},\ \mathcal{J}_{G,C^{[1]}_{n+N_C}} - \mathbf{J}_{G,P^{[1]}_{n+N_C}}\mathcal{T}_G\right]\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G\\ \vdots & \vdots\\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix}^{-1}\begin{bmatrix}I_p\\0_p\\ \vdots\\ 0_p\end{bmatrix}.$$

For the second family of bi-orthogonal matrix polynomials we proceed as follows. First, recalling Definition 1.13, we write (6.10) as follows

$$\sum_{k=0}^{n-1}\left(\hat{P}^{[2]}_k(y)\right)^{\dagger}\hat{H}_k^{-1}\hat{C}^{[1]}_k(x)W_G(x) = W_G(y)K^{(pc)}_{n-1}(x,y) + \mathcal{V}_G(x,y)$$
$$-\left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1}\ \cdots\ \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix}0_{m_Cp\times N_Gp} & -\Omega^C[n]\\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\begin{bmatrix}C^{[1]}_{n-N_G}(x)\\ \vdots\\ C^{[1]}_{n+N_C-1}(x)\end{bmatrix},$$

and the corresponding spectral jets fulfill

$$\sum_{k=0}^{n-1}\left(\hat{P}^{[2]}_k(y)\right)^{\dagger}\hat{H}_k^{-1}\mathcal{J}_{G,\hat{C}^{[1]}_k W_G} = W_G(y)\mathcal{J}_{G,K^{(pc)}_{n-1}}(y) + \mathcal{J}_{G,\mathcal{V}}(y)$$
$$-\left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1}\ \cdots\ \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix}0_{m_Cp\times N_Gp} & -\Omega^C[n]\\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{G,C^{[1]}_{n-N_G}}\\ \vdots\\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}}\end{bmatrix}.$$

From (6.17)

$$\sum_{k=0}^{n-1}\left(\hat{P}^{[2]}_k(y)\right)^{\dagger}\hat{H}_k^{-1}\mathbf{J}_{G,\hat{P}^{[1]}_k W_C}\mathcal{T}_G = W_G(y)\mathcal{J}_{G,K^{(pc)}_{n-1}}(y) + \mathcal{J}_{G,\mathcal{V}}(y)$$
$$-\left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1}\ \cdots\ \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix}0_{m_Cp\times N_Gp} & -\Omega^C[n]\\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{G,C^{[1]}_{n-N_G}}\\ \vdots\\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}}\end{bmatrix}, \quad (6.19)$$

while, from (6.9), we realize that

$$\sum_{k=0}^{n-1}\left(\hat{P}^{[2]}_k(y)\right)^{\dagger}\hat{H}_k^{-1}\mathbf{J}_{G,\hat{P}^{[1]}_k W_C}\mathcal{T}_G = W_G(y)\mathbf{J}_{G,K_{n-1}}(y)\mathcal{T}_G$$
$$-\left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1}\ \cdots\ \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix}0_{m_Cp\times N_Gp} & -\Omega^C[n]\\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\begin{bmatrix}\mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G\\ \vdots\\ \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix},$$

which can be subtracted to (6.19) to get

$$W_G(y)\left(\mathcal{J}_{G,K^{(pc)}_{n-1}}(y) - \mathbf{J}_{G,K_{n-1}}(y)\mathcal{T}_G\right) + \mathcal{J}_{G,\mathcal{V}}(y)$$

$$= \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} \cdots \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix} 0_{m_Cp\times N_Gp} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots \\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix}.$$

(6.9) also implies

$$W_G(y)\mathcal{J}_{C,K_{n-1}}(y)$$

$$= \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} \cdots \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix} 0_{m_Cp\times N_Gp} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n-N_G}} \\ \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}}\end{bmatrix}. \tag{6.20}$$

Hence, we obtain

$$\left[W_G(y)\mathcal{J}_{C,K_{n-1}}(y),\ W_G(y)\left(\mathcal{J}_{G,K^{(pc)}_{n-1}}(y) - \mathbf{J}_{G,K_{n-1}}(y)\mathcal{T}_G\right) + \mathcal{J}_{G,\mathcal{V}}(y)\right]\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix}^{-1}$$

$$= \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} \cdots \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix} 0_{m_Cp\times N_Gp} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}.$$

Now, from Definition 6.8 and the fact $\omega_{n,n-N_G} = \hat{H}_n\left(H_{n-N_G}\right)^{-1}$, we get

$$\left(\check{P}^{[2]}_n(y)\right)^{\dagger} = \left[W_G(y)\mathcal{J}_{C,K_{n-1}}(y), W_G(y)\left(\mathcal{J}_{G,K^{(pc)}_{n-1}}(y) - \mathbf{J}_{G,K_{n-1}}(y)\mathcal{T}_G\right) + \mathcal{J}_{G,\mathcal{V}}(y)\right]$$

$$\times\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G\end{bmatrix}^{-1}\begin{bmatrix}H_{n-N_G} \\ 0_p \\ \vdots \\ 0_p\end{bmatrix},$$

and the result follows. ∎

Discussion for $n < N_G$.

Proposition 6.18. *We have that*

$$\omega\left[\mathcal{J}_{C,P^{[1]}}, \mathcal{J}_{G,C^{[1]}} - \mathbf{J}_{G,P^{[1]}}\mathcal{T}_G\right] = -\hat{H}\left(\hat{S}_2\right)^{-\dagger}\begin{bmatrix}0_{N_Cp} & \mathcal{B}_G\mathcal{Q}_G \\ 0_{N_Cp} & 0_{N_Gp} \\ \vdots & \vdots\end{bmatrix}.$$

Proof. From (6.7) we deduce

$$\mathcal{J}_{G,\hat{C}^{[1]}W_G} - \begin{bmatrix} (\hat{H}(\hat{S}_2)^{-\dagger})_{[N_G]}\mathcal{B}_G\mathcal{J}_{G,\chi_{[N_G]}} \\ 0 \\ \vdots \end{bmatrix} = \omega\mathcal{J}_{G,C^{[1]}}.$$

Recalling (6.17) we obtain

$$\mathcal{J}_{G,\hat{P}^{[1]}W_C}\mathcal{T}_G - \begin{bmatrix} (\hat{H}(\hat{S}_2)^{-\dagger})_{[N_G]}\mathcal{B}_G\mathcal{J}_{G,\chi_{[N_G]}} \\ 0 \\ \vdots \end{bmatrix} = \omega\mathcal{J}_{G,C^{[1]}}.$$

Therefore, using (6.3)

$$\omega\big(\mathcal{J}_{G,C^{[1]}} - \mathtt{J}_{G,P^{[1]}}\mathcal{T}_G\big) = \begin{bmatrix} -(\hat{H}(\hat{S}_2)^{-\dagger})_{[N_G]}\mathcal{B}_G\mathcal{J}_{G,\chi_{[N_G]}} \\ 0 \\ \vdots \end{bmatrix}.$$

Observe that (6.3) also implies

$$\omega\mathcal{J}_{C,P^{[1]}} = 0.$$

■

Given a block matrix A we denote by $A_{[N],[M]}$ the truncation of A obtained by taking the first N block rows and the first M first columns. Then,

Lemma 6.19. *The following relation*

$$\omega_{[N_g],[N_C+N_G]}\left[\mathcal{J}_{C,P^{[1]}}, \mathcal{J}_{G,C^{[1]}} - \mathtt{J}_{G,P^{[1]}}\mathcal{T}_G\right]_{[N_C+N_G]}\operatorname{diag}(I_{N_C}, \mathcal{R}_G) = \left[0_{N_Gp\times N_Cp}, -\hat{H}_{[N_G]}\big(\hat{S}_2\big)^{-\dagger}_{[N_G]}\right] \tag{6.21}$$

holds.

Lemma 6.20. *The matrix* $\big(\mathcal{J}_{C,P^{[1]}}\big)_{[N_C]}$ *is nonsingular.*

Proof. It follows immediately from

$$\big(\mathcal{J}_{C,P^{[1]}}\big)_{[N_C]} = \big(S_1\big)_{[N_C]}\mathcal{Q}_C,$$

and the fact that $\mathcal{Q}_C$ is nonsingular. ■

Lemma 6.21.

$$
\begin{aligned}
-\left[\mathcal{J}_{C,P^{[1]}},\mathcal{J}_{G,C^{[1]}}-\mathtt{J}_{G,P^{[1]}}\mathcal{T}_G\right]_{[N_C+N_G]}&\operatorname{diag}(I_{N_C},\mathcal{R}_G)\\
&=\left[\begin{array}{c|c} I_{[N_C]} & 0_{pN_C\times pN_G}\\ \hline \multicolumn{2}{c}{\omega_{[N_G],[N_C+N_G]}}\end{array}\right]^{-1}\begin{bmatrix}\left(\mathcal{J}_{C,P^{[1]}}\right)_{[N_C]} & 0_{pN_C\times pN_G}\\ 0_{pN_G\times pN_C} & \hat{H}_{[N_G]}\end{bmatrix}\\
&\quad\times\begin{bmatrix} I_{[N_C]} & \left(\left(\mathcal{J}_{C,P^{[1]}}\right)_{[N_C]}\right)^{-1}\left(\mathcal{J}_{G,C^{[1]}}-\mathtt{J}_{G,P^{[1]}}\mathcal{T}_G\right)_{[N_C],[N_G]}\mathcal{R}_G\\ 0_{N_Gp\times N_Cp} & (\hat{S}_2)^{-\dagger}_{[N_G]}\end{bmatrix}
\end{aligned}
$$

holds.

Proof. (6.21) can be written as

$$
\begin{aligned}
&-\left[\begin{array}{c|c} I_{[N_C]} & 0_{pN_C\times pN_G}\\ \hline \multicolumn{2}{c}{\omega_{[N_G],[N_C+N_G]}}\end{array}\right]\left[\mathcal{J}_{C,P^{[1]}},\mathcal{J}_{G,C^{[1]}}-\mathtt{J}_{G,P^{[1]}}\mathcal{T}_G\right]_{[N_C+N_G]}\operatorname{diag}(I_{N_C},\mathcal{R}_G)\\
&=\begin{bmatrix}\left(\mathcal{J}_{C,P^{[1]}}\right)_{[N_C]} & 0_{pN_C\times pN_G}\\ 0_{pN_G\times pN_C} & \hat{H}_{[N_G]}\end{bmatrix}\begin{bmatrix} I_{[N_C]} & \left(\left(\mathcal{J}_{C,P^{[1]}}\right)_{[N_C]}\right)^{-1}\left(\mathcal{J}_{G,C^{[1]}}-\mathtt{J}_{G,P^{[1]}}\mathcal{T}_G\right)_{[N_C],[N_G]}\mathcal{R}_G\\ 0_{N_Gp\times N_Cp} & \left((\hat{S}_2)_{[N_G]}\right)^{-\dagger}\end{bmatrix},
\end{aligned}
$$

and the result follows. ∎

Lemma 6.22. *For $n\in\{1,\dots,N_G-1\}$ we find*

$$
\begin{aligned}
\hat{H}_n&=-\Theta_*\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_0} & \left(\mathcal{J}_{C^{[1]}_0}-\mathtt{J}_{P^{[1]}_0}\mathcal{T}\right)\mathcal{R}_{G,n}\\ \vdots & \vdots\\ \mathcal{J}_{C,P^{[1]}_{N_C+n-1}} & \left(\mathcal{J}_{C^{[1]}_{n-1}}-\mathtt{J}_{P^{[1]}_{N_C+n-1}}\mathcal{T}\right)\mathcal{R}_{G,n}\end{bmatrix},\\
\omega_{n,k}&=\Theta_*\left[\begin{array}{cc|c}\mathcal{J}_{C,P^{[1]}_0} & \left(\mathcal{J}_{C^{[1]}_0}-\mathtt{J}_{P^{[1]}_0}\mathcal{T}\right)\mathcal{R}_{G,n} & \\ \vdots & \vdots & e_k\\ \mathcal{J}_{C,P^{[1]}_{N_C+n}} & \left(\mathcal{J}_{C^{[1]}_{N_C+n}}-\mathtt{J}_{P^{[1]}_{N_C+n}}\mathcal{T}\right)\mathcal{R}_{G,n} & \end{array}\right], && 0\leq k<N_C+n,\\
\left((\hat{S}_2)^{\dagger}\right)_{n,k}&=\Theta_*\begin{bmatrix}\mathcal{J}_{C,P^{[1]}_0} & \left(\mathcal{J}_{C^{[1]}_0}-\mathtt{J}_{P^{[1]}_0}\mathcal{T}\right)\mathcal{R}_{G,n+1}\\ \vdots & \vdots\\ \mathcal{J}_{C,P^{[1]}_{N_C+n-1}} & \left(\mathcal{J}_{C^{[1]}_{n-1}}-\mathtt{J}_{P^{[1]}_{N_C+n-1}}\mathcal{T}\right)\mathcal{R}_{G,n+1}\\ \multicolumn{2}{c}{(e_{N_C+k})^{\dagger}}\end{bmatrix}, && 0\leq k<N_G.
\end{aligned}
$$

Here we have used the matrices $e_k=\left[0_p,\dots,0_p,I_p,0_p,\dots,0_p)\right]^{\dagger}\in\mathbb{C}^{(N_C+n+1)p\times p}$ with all its $p\times p$ blocks the zero matrix 0_p, but the k-th block is the identity matrix I_p.

Theorem 6.23 (Spectral Christoffel–Geronimus-Uvarov formulas)**.** *For $n<N$ and monic polynomial linear spectral perturbations , we have the following last quasi-determinant expressions for*

the perturbed bi-orthogonal matrix polynomials

$$\hat{P}_n^{[1]}(x)W_C(x) = \Theta_* \begin{bmatrix} \mathcal{J}_{C,P_0^{[1]}} & \left(\mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T}\right)\mathcal{R}_{G,n} & P_n^{[1]}(x) \\ \vdots & \vdots & \vdots \\ \mathcal{J}_{C,P_{N_C+n}^{[1]}} & \left(\mathcal{J}_{C_{N_C+n}^{[1]}} - \mathbf{J}_{P_{N_C+n}^{[1]}}\mathcal{T}\right)\mathcal{R}_{G,n} & P_{n+N_C+n}^{[1]}(x) \end{bmatrix},$$

$$\left(\hat{P}_n^{[2]}(x)\right)^\dagger = \Theta_* \begin{bmatrix} \mathcal{J}_{C,P_0^{[1]}} & \left(\mathcal{J}_{C_0^{[1]}} - \mathbf{J}_{P_0^{[1]}}\mathcal{T}\right)\mathcal{R}_{G,n+1} \\ \vdots & \vdots \\ \mathcal{J}_{C,P_{N_C+n-1}^{[1]}} & \left(\mathcal{J}_{C_{n-1}^{[1]}} - \mathbf{J}_{P_{N_C+n-1}^{[1]}}\mathcal{T}\right)\mathcal{R}_{G,n+1} \\ 0_{p\times pN_C} & \left(\chi_{[N_G]}(x)\right)^\dagger \end{bmatrix}.$$

Proof. From (6.3) and (1.8), for $n < N_G$ we have

$$\hat{P}_n^{[1]}(x)W_C(x) = P_{N_C+n}^{[1]}(x) + \left[\omega_{n,0},\dots,\omega_{n,N_C+n-1}\right]\begin{bmatrix} P_0^{[1]}(x) \\ \vdots \\ P_{N_C+n-1}^{[1]}(x) \end{bmatrix},$$

$$\left(\hat{P}_n^{[2]}\right)^\dagger(x) = I_p x^n + \left[I_p,\dots,I_p x^{n-1}\right]\begin{bmatrix} \left((\hat{S}_2)^\dagger\right)_{0,n} \\ \vdots \\ \left((\hat{S}_2)^\dagger\right)_{n-1,n} \end{bmatrix}.$$

Then, Lemma 5.22 gives the stated result. ∎

6.4 Mixed spectral/nonspectral Christoffel–Geronimus–Uvarov formulas

Recall that we consider the linear spectral transformation as a composition of a Geronimus transformation (in the first step), applying nonspectral techniques, and then a Christoffel transformation, for which we will apply spectral techniques. In this situation, we still require the leading coefficient A_{C,N_C} must be nonsingular, but we give a freedom degree for $W_G(x)$ of such a condition. Thus, we consider

$$\check{u} := u(W_G(x))^{-1} + v_G, \qquad v_G := \sum_{a=1}^{q_G}\sum_{j=1}^{s_{G,a}}\sum_{m=0}^{\kappa_{G,j}^{(a)}-1}(-1)^m\delta^{(m)}(x-x_{G,a})\frac{\xi_{j,m}^{[a]}}{m!}l_{G,j}^{(a)}(x),$$

so that, after a Christoffel transformation, the linear spectral transformation is achieved

$$\hat{u} = W_C(x)\check{u}.$$

As for the Geronimus case, we consider

Definition 6.24. *For a given perturbed matrix of functionals $\check{u}$ we define a semi-infinite block matrix*

$$R := \left\langle P^{[1]}, \chi(x) \right\rangle_{\check{u}} = \left\langle P^{[1]}, \chi(x) \right\rangle_{uW_G^{-1}} + \left\langle P^{[1]}, \chi(x) \right\rangle_{v_G}.$$

Proposition 6.25. *The equations*

$$R = S_1 \check{M}, \tag{6.22}$$

$$\omega R = \hat{H}\left(\hat{S}_2\right)^{-\dagger} \tag{6.23}$$

are fulfilled.

Proof. (6.22) is a consequence of its definition. Let us show (6.23)

$$\begin{aligned} \omega R &= \hat{S}_1 W_C(\Lambda)\left(S_1\right)^{-1} S_1 \check{M} \\ &= \hat{S}_1 W_C(\Lambda)\check{M} \\ &= \hat{S}_1 \hat{M} \\ &= \hat{H}\left(\hat{S}_2\right)^{-\dagger}. \end{aligned}$$

■

Thus, we conclude

Proposition 6.26. *The matrix R satisfies*

$$(\omega R)_{n,l} = \begin{cases} 0_p, & l \in \{0,\dots,n-1\}, \\ \hat{H}_n, & n = l. \end{cases}$$

Moreover, the matrix $\begin{bmatrix} R_{0,0} & \dots & R_{0,n-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \dots & R_{n-1,n-1} \end{bmatrix}$ *is nonsingular.*

As for the pure Geronimus situation we consider

Definition 6.27. *We introduce the matrix polynomials $R^K_{n,l}(y) \in \mathbb{C}^{p\times p}[y]$, $l \in \{0,\dots,n-1\}$, given by*

$$\begin{aligned} R^K_{n,l}(y) :&= \left\langle W_G(y)K_{n-1}(x,y), I_p x^l \right\rangle_{\check{u}} - I_p y^l \\ &= \left\langle W_G(y)K_{n-1}(x,y), I_p x^l \right\rangle_{uW_G^{-1}} + \left\langle W_G(y)K_{n-1}(x,y), I_p x^l \right\rangle_{v_G} - I_p y^l. \end{aligned}$$

Proposition 6.28. *For $l \in \{0,1,\dots,n-1\}$ and $n \geq N_G$, we get*

$$R^K_{n,l}(y) = \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^\dagger \left(\hat{H}_{n-m_C}\right)^{-1}, \dots, \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^\dagger \left(\hat{H}_{n+N_G-1}\right)^{-1} \right] \\ \times \begin{bmatrix} 0_{m_C p\times N_G p} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_G p\times N_C p} \end{bmatrix} \begin{bmatrix} R_{n-N_G,l} \\ \vdots \\ R_{n+N_C-1,l} \end{bmatrix}. \tag{6.24}$$

Proof. It follows from (6.9), Definition 6.24, and (1.15). ∎

Definition 6.29. *For $n \geq N_G$, let us assume that the matrix*

$$\Phi_n := \begin{bmatrix} \mathcal{I}_{C,P^{[1]}_{n-N_G}} & R_{n-N_G,0} & \dots & R_{n-N_G,n-1} \\ \vdots & \vdots & & \vdots \\ \mathcal{I}_{C,P^{[1]}_{n+N_C-1}} & R_{n+N_C-1,0} & \dots & R_{n+N_C-1,n-1} \end{bmatrix} \in \mathbb{C}^{(N_C+N_G)p\times(N_C+n)p}$$

has full rank, i.e. $\operatorname{rank}(\Phi_n) = N_C + N_G$. *Then, there exist nonsingular square submatrices $\Phi_n^\square \in \mathbb{C}^{(N_C+N_G)p\times(N_C+N_G)p}$ of Φ_n, and we will refer to them as poised submatrices. We also consider*

$$\varphi_n := \left[\mathcal{I}_{C,P^{[1]}_{n+N_C}}, R_{n+N_C,0}, \dots, R_{n+N_C,n-1}\right] \in \mathbb{C}^{p\times(N_C+n)p}$$

and the submatrices $\varphi_n^\square \in \mathbb{C}^{p\times(N_C+N_G)p}$ corresponding to the selection of columns to built the poised submatrix $\Phi_n^\square$. We also consider

$$\varphi_n^K(y) = \left[W_G(y)\mathcal{I}_{C,K_{n-1}}(y), R^K_{n,0}(y), \dots, R^K_{n,n-1}(y)\right] \in \mathbb{C}^{p\times(N_C+n)p}[y]$$

and $(\varphi_n^K(y))^\square$.

Proposition 6.30. *If $A_{G,N}$ is nonsingular, then*

$$\begin{bmatrix} \mathcal{I}_{C,P^{[1]}_{n-N_G}} & R_{n-N_G,0} & \dots & R_{n-N_G,N_G-1} \\ \vdots & \vdots & & \vdots \\ \mathcal{I}_{C,P^{[1]}_{n+N_C-1}} & R_{n+N_C-1,0} & \dots & R_{n+N_C-1,N_G-1} \end{bmatrix}$$

is nonsingular.

Proof. From Proposition 6.26 we get the system

$$\left[\omega_{n,n-N_G}, \dots, \omega_{n,n+N_C-1}\right] \begin{bmatrix} R_{n-N_G,l} \\ \vdots \\ R_{n+N_C-1,l} \end{bmatrix} = -R_{n+N_C,l},$$

for $l \in \{0,1,\dots,n-1\}$. In particular, the resolvent vector $\left[\omega_{n,n-N}, \dots, \omega_{n,n-1}\right]$ is a solution of the linear system

$$\left[\omega_{n,n-N_G}, \dots, \omega_{n,n+N_C-1}\right] \begin{bmatrix} \mathcal{I}_{C,P^{[1]}_{n-N_G}} & R_{n-N_G,0} & \dots & R_{n-N_G,N_G-1} \\ \vdots & \vdots & & \vdots \\ \mathcal{I}_{C,P^{[1]}_{n+N_C-1}} & R_{n+N_C-1,0} & \dots & R_{n+N_C-1,N_G-1} \end{bmatrix}$$

$$= -\left[\mathcal{J}_{C,P^{[1]}_{n+N_C}}, R_{n+N_C,0}, \dots, R_{n+N_C,N_G-1}\right]. \quad (6.25)$$

Let us discuss the uniqueness of the solutions of this system of linear linear. Assume that there exists another solution $\left[\tilde{\omega}_{n,n-N_G}, \dots, \tilde{\omega}_{n,n+N_C-1}\right]$. Then, consider the monic matrix polynomial

$$Q_{n+N_C}(x) = P^{[1]}_{n+N_C}(x) + \tilde{\omega}_{n,n+N_C-1}P^{[1]}_{n+N_C-1}(x) + \dots + \tilde{\omega}_{n,n-N_G}P^{[1]}_{n-N_G}(x).$$

Because $\left[\tilde{\omega}_{n,n-N}, \dots, \tilde{\omega}_{n,n-1}\right]$ solves (6.25) we get

$$\mathcal{J}_{C,Q_{n+N_C}} = 0_{p\times N_C p}, \quad (6.26)$$

$$\langle Q_{n+N_C}(x), I_p x^l\rangle_{\check{u}} = 0_p, \qquad l \in \{0, \dots, N_G - 1\}. \quad (6.27)$$

Using Lemma 1.37, (6.26) can be expressed as follows

$$\mathcal{J}_{C,Q_{n+N_C}} = \left[\left(Q_{n+N_C}\right)_0, \dots, \left(Q_{n+N_C}\right)_{n+N_C}\right]\begin{bmatrix} X_C \\ XJ_C \\ \vdots \\ X_C(J_C)^{n+N_C}\end{bmatrix} = 0_{p,N_C p},$$

where (X_C, J_C) is a Jordan pair for the perturbing polynomial $W_C(x)$. But, from Proposition 1.32, this is a necessary and sufficient condition for $W_C(x)$ to be a right divisor of the polynomial $Q_{n+N_C}(x)$, so that we can write

$$Q_{n+N_C}(x) = \tilde{P}_n(x)W_C(x),$$

where $\tilde{P}_n(x)$ is a monic polynomial of degree n. Then, (6.27) reads

$$\langle \tilde{P}_n(x), I_p x^l\rangle_{\hat{u}} = 0_p, \qquad l \in \{0, \dots, N_G - 1\}.$$

We now proceed as we did in Proposition 5.33. We first notice that Lemma 5.32 can be applied again to get

$$\langle P^{[1]}_m(x), I_p x^l\rangle_{\check{u}} = \langle P^{[1]}_m(x), \beta_l(x)\rangle_{\check{u}}$$

for $l < m + N_G$. Hence, when $l \in \{0, \dots, n-1\}$ we have

$$\begin{aligned}
\langle Q_{n+N_C}(x), I_p x^l\rangle_{\check{u}} = \langle \tilde{P}_n(x), I_p x^l\rangle_{\hat{u}} &= \sum_{k=n-N_G}^{N_C+n} \omega_{n,k}\langle P^{[1]}_k(x), I_p x^l\rangle_{\check{u}} \\
&= \sum_{k=n-N_G}^{N_C+n} \omega_{n,k}\langle P^{[1]}_k(x), \alpha_l(x)\rangle_u + \langle \tilde{P}_n(x), \beta_l(x)\rangle_{\check{u}} \\
&= 0_p.
\end{aligned}$$

Now, the uniqueness for bi-orthogonal polynomial families implies

$$\tilde{P}_n(x) = \hat{P}_n^{[1]}(x),$$

and, considering (6.3), we infer that there is an unique solution of (6.25). Thus,

$$\begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & R_{n-N_G,0} & \dots & R_{n-N_G,N_G-1} \\ \vdots & \vdots & & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & R_{n+N_C-1,0} & \dots & R_{n+N_C-1,N_G-1} \end{bmatrix}$$

is nonsingular and, therefore, is a poised submatrix. ■

Proposition 6.31. *For $n \geq N_G$ and a full rank matrix Φ_n, let us take a poised submatrix $\Phi_n^{\square}$. Then,*

$$\left[\omega_{n,n-N_G}, \dots, \omega_{n,n+N_C-1}\right] = -\varphi_n^{\square}\left(\Phi_n^{\square}\right)^{-1}.$$

Proof. From Proposition 6.26, for $n \geq N_G$ we get

$$\left[\omega_{n,n-N_G}, \dots, \omega_{n,n+N_C-1}\right] \begin{bmatrix} R_{n-N_G,0} & \dots & R_{n-N_G,n-1} \\ \vdots & & \vdots \\ R_{n+N_C-1,0} & \dots & R_{n+N_C-1,n-1} \end{bmatrix} = -\left[R_{n+N_C,0}, \dots, R_{n+N_C,n-1}\right].$$

Using (6.18) we can extend this equation to

$$\left[\omega_{n,n-N_G}, \dots, \omega_{n,n+N_C-1}\right]\Phi_n = -\varphi_n,$$

and the result follows. ■

Theorem 6.32 (Mixed spectral/nonspectral matrix Christoffel–Geronimus–Uvarov formulas)**.** *Given a matrix Geronimus-Uvarov transformation, the corresponding perturbed polynomials can be expressed, for $n \geq N_G$, as follows*

$$\hat{P}_n^{[1]}(x)W_C(x) = \Theta_* \left[\begin{array}{c|c} \Phi_n^{\square} & \begin{matrix} P^{[1]}_{n-N_G}(x) \\ \vdots \\ P^{[1]}_{n+N_C-1}(x) \end{matrix} \\ \hline \varphi_n^{\square} & P^{[1]}_{n+N_C}(x) \end{array}\right], \quad \left(\hat{P}_n^{[2]}(x)\right)^{\dagger} A_{G,N_G} = -\Theta_* \left[\begin{array}{c|c} \Phi_n^{\square} & \begin{matrix} H_{n-N_G} \\ 0_p \\ \vdots \\ 0_p \end{matrix} \\ \hline \left(\varphi_n^K(x)\right)^{\square} & 0_p \end{array}\right].$$

The corresponding quasitau matrices are

$$\hat{H}_n = \Theta_* \left[\begin{array}{c|c} \Phi_n^{\square} & \begin{matrix} R_{n-N_G,n} \\ \vdots \\ R_{n+N_C-1,n} \end{matrix} \\ \hline \varphi_n^{\square} & R_{n+N_C,n} \end{array}\right].$$

Proof. From the connection formula (6.3) we find

$$\hat{P}_n^{[1]}(x)W_C(x) = [\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}]\begin{bmatrix} P^{[1]}_{n-N_G}(x) \\ \vdots \\ P^{[1]}_{n+N_C-1}(x)\end{bmatrix} + P^{[1]}_{n+N_C}(x),$$

and from Proposition 6.26 we deduce

$$\hat{H}_n = [\omega_{n,n-N_G},\dots,\omega_{n,n+N_C-1}]\begin{bmatrix} R_{n-N_G,n} \\ \vdots \\ R_{n+N_C-1,n}\end{bmatrix} + R_{n+N_C,n}.$$

Then, the result follows from Proposition 6.31. From (6.24) and (6.20) we get

$$\varphi_n^K(y) = \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} \quad \cdots \quad \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix} 0_{m_Cp\times N_Gp} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}\Phi_n,$$

so that

$$\left(\varphi_n^K(y)\right)^{\square}\left(\Phi_n^{\square}\right)^{-1} = \left[\left(\hat{P}^{[2]}_{n-m_C}(y)\right)^{\dagger}(\hat{H}_{n-m_C})^{-1} \cdots \left(\hat{P}^{[2]}_{n+N_G-1}(y)\right)^{\dagger}(\hat{H}_{n+N_G-1})^{-1}\right]\begin{bmatrix} 0_{m_Cp\times N_Gp} & -\Omega^C[n] \\ \Omega^G[n] & 0_{N_Gp\times N_Cp}\end{bmatrix}.$$

In particular, recalling (6.5) we deduce that

$$(\check{P}_n^{[2]}(y))^{\dagger}A_{G,N_G} = \left(\varphi_n^K(y)\right)^{\square}\left(\Phi_n^{\square}\right)^{-1}\begin{bmatrix} H_{n-N} \\ 0_p \\ \vdots \\ 0_p\end{bmatrix},$$

and the expression for the perturbation of the second family of bi-orthogonal polynomials follows. ■

Discussion for $n < N_G$

We proceed similarly as we did in Subsection 6.3. Let us notice that, from (6.3) and (6.22), we get

$$\omega\left[\mathcal{J}_{C,P^{[1]}},R\right] = \hat{H}\left(\hat{S}_2\right)^{-\dagger}[0,I].$$

Therefore, we conclude

$$\omega_{[N_G,N_C+N_G]}\left[\mathcal{J}_{C,P^{[1]}},R\right]_{[N_C+N_G]} = \left[0_{N_Gp\times N_Cp}, -\hat{H}_{[N_G]}\left(\hat{S}_2\right)^{-\dagger}_{[N_G]}\right].$$

Given a block matrix A we denote by $A_{[N,M]}$ the truncation obtained by taking the first N block rows and the first M first block columns. Then, we easily conclude

Lemma 6.33. *The following Gauss-Borel factorization is fulfilled*

$$\left[\mathcal{J}_{C,P^{[1]}},R\right]_{[N_C+N_G]}=\left[\begin{array}{c|c}I_{[N_C]} & 0_{pN_C\times pN_G}\\ \hline \multicolumn{2}{c}{\omega_{[N_G],[N_C+N_G]}}\end{array}\right]^{-1}\begin{bmatrix}\left(\mathcal{J}_{C,P^{[1]}}\right)_{[N_C]} & 0_{pN_C\times pN_G}\\ 0_{pN_G\times pN_C} & \hat{H}_{[N_G]}\end{bmatrix}\begin{bmatrix}I_{[N_C]} & R_{[N_C,N_G]}\\ 0_{N_Gp\times N_Cp} & \left(\hat{S}_2\right)^{-\dagger}_{[N_G]}\end{bmatrix}.$$

Therefore, for $n\in\{1,\dots,N_G-1\}$ we have

$$\hat{H}_n=\Theta_*\begin{bmatrix}\mathcal{J}_{C,P_0^{[1]}} & R_{0,0} & \dots & R_{0,n-1}\\ \vdots & \vdots & & \vdots\\ \mathcal{J}_{C,P_{N_C+n-1}^{[1]}} & R_{N_C+n-1,0} & \dots & R_{N_C+n-1,n-1}\end{bmatrix},$$

$$\omega_{n,k}=\Theta_*\left[\begin{array}{cccc|c}\mathcal{J}_{C,P_0^{[1]}} & R_{0,0} & \dots & R_{0,n-1} & \\ \vdots & \vdots & & & e_k\\ \mathcal{J}_{C,P_{N_C+n}^{[1]}} & R_{N_C+n,0} & \dots & R_{N_C+n,n-1} & \end{array}\right],\qquad 0\leq k<N_C+n,$$

$$\left((\hat{S}_2)^{\dagger}\right)_{n,k}=\Theta_*\begin{bmatrix}\mathcal{J}_{C,P_0^{[1]}} & R_{0,0} & \dots & R_{0,n}\\ \vdots & \vdots & & \\ \mathcal{J}_{C,P_{N_C+n-1}^{[1]}} & R_{N_C+n,0} & \dots & R_{N_C+n,n}\\ & (e_{N_C+k})^{\dagger} & & \end{bmatrix},\qquad 0\leq k<N_G.$$

Theorem 6.34 (Mixed spectral/nonspectral Christoffel–Geronimus–Uvarov formulas)**.** *For $n<N_G$, the perturbed bi-orthogonal matrix polynomials have the following quasideterminatal expressions*

$$\hat{P}_n^{[1]}(x)W_C(x)=\Theta_*\begin{bmatrix}\mathcal{J}_{C,P_0^{[1]}} & R_{0,0} & \dots & R_{0,n-1} & P_n^{[1]}(x)\\ \vdots & \vdots & & \vdots & \vdots\\ \mathcal{J}_{C,P_{N_C+n}^{[1]}} & R_{N_C+n,0} & \dots & R_{N_C+n,n-1} & P_{n+N_C}^{[1]}(x)\end{bmatrix},$$

$$\left(\hat{P}_n^{[2]}(x)\right)^{\dagger}=\Theta_*\begin{bmatrix}\mathcal{J}_{C,P_0^{[1]}} & R_{0,0} & \dots & R_{0,n}\\ \vdots & & & \vdots\\ \mathcal{J}_{C,P_{N_C+n-1}^{[1]}} & R_{N_C+n-1,0} & \dots & R_{N_C+n-1,n}\\ 0_{p\times pN_C} & & \left(\chi(x)_{[N_G]}\right)^{\dagger} & \end{bmatrix}.$$

6.5 Spectral versus nonspectral

Proposition 6.35. *For matrix Geronimus-Uvarov transformations*

$$\mathcal{J}_{G,C^{[1]}}-\mathbf{J}_{G,P^{[1]}}\mathcal{T}=-R^{(N_G)}\mathcal{B}_G\mathcal{Q}_G.$$

Proof. From (6.7) we deduce that

$$\hat{C}^{[1]}(x)W_G(x) - \hat{H}\left(\hat{S}_2\right)^{-\dagger}\begin{bmatrix} \mathcal{B}_G(\chi(x))_{[N_G]} \\ 0_p \\ \vdots \end{bmatrix} = \omega C^{[1]}(x),$$

so that

$$\mathcal{J}_{G,\hat{C}^{[1]}W_G} - \hat{H}\left(\hat{S}_2\right)^{-\dagger}\begin{bmatrix} \mathcal{B}_G Q_G \\ 0_p \\ \vdots \end{bmatrix} = \omega \mathcal{J}_{G,C^{[1]}}.$$

By using (6.17), (6.3), and (6.23) we get

$$\omega\left(\mathcal{J}_{G,C^{[1]}} - \mathbf{J}_{G,P^{[1]}}\mathcal{T}_G\right) = -\omega R\begin{bmatrix} \mathcal{B}_G Q_G \\ 0_p \\ \vdots \end{bmatrix}.$$

From (6.23) and the fact that ω is nonsingular, the result follows. ∎

Corollary 6.36. *If $A_{N,G}$ is nonsingular, then the matrix*

$$\begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G \end{bmatrix}$$

is nonsingular.

Proof. Proposition 6.35 implies

$$\begin{bmatrix} \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots \\ \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G \end{bmatrix} = -\begin{bmatrix} R_{n-N_G,0} & & R_{n-N_G,N_G-1} \\ \vdots & & \vdots \\ R_{n+N_C-1,0} & \dots & R_{n+N_C-1,N_G-1} \end{bmatrix}\mathcal{R}_G^{-1}.$$

Thus,

$$\begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & \mathcal{J}_{G,C^{[1]}_{n-N_G}} - \mathbf{J}_{G,P^{[1]}_{n-N_G}}\mathcal{T}_G \\ \vdots & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & \mathcal{J}_{G,C^{[1]}_{n+N_C-1}} - \mathbf{J}_{G,P^{[1]}_{n+N_C-1}}\mathcal{T}_G \end{bmatrix}$$

$$= \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & R_{n-N_G,0} & \dots & R_{n-N_G,N_G-1} \\ \vdots & \vdots & & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & R_{n+N_C-1,0} & \dots & R_{n+N_C-1,N_G-1} \end{bmatrix} \operatorname{diag}(I_{N_G p}, -\mathcal{R}_G^{-1}).$$

Finally, Proposition 6.30 and Lemma 1.40 give the result. ■

6.6 Applications

6.6.1 Matrix Geronimus-Uvarov transformation and Christoffel transformations with singular leading coefficients

In Chapter 4 we deal with Christoffel transformations for matrix orthogonal polynomials. However, for this discussion it was necessary to assume that the leading coefficient of the polynomial involved in the perturbation was nonsingular, so that we could apply all the spectral machinery described in [77]. In Subsection 4.3, we considered unimodular Christoffel transformations, which have been broadly studied inside the matrix orthogonal polynomials community. As it was shown therein, despite the nonspectral condition, the Geronimus transformation can provide interesting formulas for the perturbed matrix orthogonal polynomials. We extend now these considerations for a Christoffel transformation with singular leading coefficient and not necessarily unimodular. The idea is to use the adjugate or classical adjoint of a matrix polynomial $W(x)$, $\operatorname{adj}(W(x))$, defined as the transpose of the matrix of cofactors, also known as the classic adjoint (see [83]). From the Laplace formula

$$W(x)\operatorname{adj}(W(x)) = \operatorname{adj}(W(x))W(x) = \det(W(x))I_p,$$

As $\det W(x)$ is a scalar polynomial with $\deg\det W(x) \leq Np$, the degree of the adjugate polynomial $\operatorname{adj}(W(x))$ is bounded as follows

$$\deg\operatorname{adj}(W(x)) \leq N(p-1).$$

We point out the relations

$$\begin{aligned} \big(W(x)\big)^{-1} &= \frac{1}{\det(W(x))}\operatorname{adj}(W(x)), \\ W(x) &= \det(W(x))\big(\operatorname{adj}(W(x))\big)^{-1}, \end{aligned}$$

which will be instrumental in the sequel. We study the Christoffel transformation[2]

$$\hat{u} = uW(x),$$

[2]Notice that the matrix Christoffel transformation $\hat{u} = W(x)u$ is a transposition of this one.

as the following massless matrix Geronimus-Uvarov transformation

$$\hat{u} = W_C(x)u\big(W_G(x)\big)^{-1}, \qquad W_C(x) := I_p \det(W(x)), \qquad W_G(x) := \operatorname{adj}(W(x)),$$

with $N_C = \deg W_C(x) \leq Np$ and $N_G = \deg W_G(x)$, where $N := \deg W(x)$ and $N_C \leq N_G + N$. Therefore, we could apply our results for matrix Geronimus-Uvarov transformations. To cover the case when the matrix $W_G(x)$ has a singular leading coefficient, we will use the mixed spectral-nonspectral Christoffel–Geronimus–Uvarov formula of Theorem 6.32. That is, the perturbed polynomials can be expressed, for $n \geq N_G$, as follows

$$\hat{P}_n^{[1]}(x)\det(W(x)) = \Theta_* \left[\begin{array}{c|c} \Phi_n^{\square} & \begin{matrix} P^{[1]}_{n-N_G}(x) \\ \vdots \\ P^{[1]}_{n+N_C-1}(x) \end{matrix} \\ \hline \varphi_n^{\square} & P^{[1]}_{n+N_C}(x) \end{array}\right], \quad \big(\hat{P}_n^{[2]}(x)\big)^{\dagger} A_N = -\Theta_* \left[\begin{array}{c|c} \Phi_n^{\square} & \begin{matrix} H_{n-N_G} \\ 0_p \\ \vdots \\ 0_p \end{matrix} \\ \hline \big(\varphi_n^K(x)\big)^{\square} & 0_p \end{array}\right],$$

and the corresponding quasitau matrices are

$$\hat{H}_n = \Theta_* \left[\begin{array}{c|c} \Phi_n^{\square} & \begin{matrix} R_{n-N_G,n} \\ \vdots \\ R_{n+N_C-1,n} \end{matrix} \\ \hline \varphi_n^{\square} & R_{n+N_C,n} \end{array}\right].$$

The reader should remind at this point Definition 6.29. Recall that $\Phi_n^{\square} \in \mathbb{C}^{(N_C+N_G)p\times(N_C+N_G)p}$ is a nonsingular submatrix of

$$\Phi_n := \begin{bmatrix} \mathcal{J}_{C,P^{[1]}_{n-N_G}} & R_{n-N_G,0} & \dots & R_{n-N_G,n-1} \\ \vdots & \vdots & & \vdots \\ \mathcal{J}_{C,P^{[1]}_{n+N_C-1}} & R_{n+N_C,0} & \dots & R_{n+N_C,n-1} \end{bmatrix} \in \mathbb{C}^{(N_C+N_G)p\times(N_C+n)p}$$

as well as $\varphi_n^{\square}, (\varphi_n^K)^{\square} \in \mathbb{C}^{p\times(N_C+N_G)p}$ correspond to the same choice of columns of

$$\begin{aligned} \varphi_n &= \Big[\mathcal{J}_{C,P^{[1]}_{n+N_C}}, R_{n+N_C,0}, \dots, R_{n+N_C,n-1}\Big] \in \mathbb{C}^{p\times(N_C+n)p}, \\ \varphi_n^K &= \big[W_G(y)\mathcal{J}_{C,K_{n-1}}(y), R^K_{n,0}(y), \dots, R^K_{n,n-1}(y)\big] \in \mathbb{C}^{p\times(N_C+n)p}[y], \end{aligned}$$

respectively. Now, the entries of R are

$$R_{n,m} = \left\langle P_n^{[1]}(x), x^m \frac{W(x)}{\det(W(x))} \right\rangle_u$$

while the root spectral jet $\mathcal{J}_{C,P}$ is easy to compute since $W_C(x) = \det(W(x))$. In this case,

$$\mathcal{J}_{C,P} = \left[P(x_1), P'(x_1), \dots, \frac{P^{(\alpha_1-1)}(x_1)}{\alpha_1!}, \dots, P(x_q), P'(x_q), \dots, \frac{P^{(\alpha_q-1)}(x_q)}{\alpha_q!}\right],$$

where x_a are the eigenvalues, with corresponding multiplicities α_a, of $W(x)$. Thus, we deduce the explicit expression

$$\Phi = \begin{bmatrix} P^{[1]}_{n-N_G}(x_1) & \dots & \dfrac{(P^{[1]}_{n-N_G})^{(\alpha_q-1)}(x_q)}{\alpha_q!} & \left\langle P^{[1]}_{n-N_G}(x), \dfrac{W(x)}{\det(W(x))} \right\rangle_u & \dots & \left\langle P^{[1]}_{n-N_G}(x), \dfrac{x^{n-1}W(x)}{\det(W(x))} \right\rangle_u \\ \vdots & & \vdots & \vdots & & \vdots \\ P^{[1]}_{n+N_C-1}(x_1) & \dots & \dfrac{(P^{[1]}_{n+N_C-1})^{(\alpha_q-1)}(x_q)}{\alpha_q!} & \left\langle P^{[1]}_{n+N_C-1}(x), \dfrac{W(x)}{\det(W(x))} \right\rangle_u & \dots & \left\langle P^{[1]}_{n+N_C-1}(x), \dfrac{x^{n-1}W(x)}{\det(W(x))} \right\rangle_u \end{bmatrix}$$

from where a poised submatrix, which exists for an appropriate selection of columns, must be picked.

6.6.2 Spectral symmetric transformations

If the matrix of linear functionals u is symmetric, $u = u^\dagger$, the bi-orthogonality becomes an orthogonality condition, and so we have orthogonal matrix polynomials. The perturbations we have considered so far do not respect this condition. The study of transformations preserving this symmetric condition yields transformations

$$\hat{u} = W(x)u(W(x))^\dagger,$$

which we call of the symmetric Christoffel type, or (we omit the masses for sake of simplicity)

$$\hat{u} = (W(x))^{-1}u(W(x))^{-\dagger},$$

and we call it the massless symmetric Geronimus type transformations. They can be understood using the adjugate technique at the light of the matrix Geronimus-Uvarov transformation techniques we have discussed previously. We need to assume that the leading coefficient of $W(x)$ is nonsingular and, hence, spectral techniques could be applied. For example, the symmetric Christoffel transformation can be written as the following matrix Geronimus-Uvarov transformation

$$\hat{u} = W_C(x)u(W_G(x))^{-1}, \qquad W_C := \det(W(x))W(x), \qquad W_G := (\operatorname{adj}(W(x)))^\dagger,$$

with polynomial degrees $N_C = N(p+1)$ and $N_G = N(p-1)$. The massless symmetric Geronimus transformation can be understood as the following matrix Geronimus-Uvarov transformation

$$\hat{u} = W_C(x)u(W_G(x))^{-1}, \qquad W_C := \operatorname{adj}(W(x)), \qquad W_G := \det(W(x))(W(x))^\dagger.$$

Now, the degrees are $N_C = N(p-1)$, $N_G = N(p+1)$, respectively. Observe that for the symmetric Christoffel transformation we have $\det W_C(x) = (\det W(x))^{p+1}$, so that eigenvalues coincide but multiplicities are multiplied by $p+1$. The same happens for the massless symmetric Geronimus transformations and $W_G(x)$. Notice also that as $\det(\operatorname{adj}(W(x))) = \det(W(x))^{p-1}$ we see that the eigenvalues of $W_G(x)$ in the symmetric Christoffel and those of $W_C(x)$ in the symmetric Geronimus are those of $W(x)$ but with multiplicities multiplied by $p-1$. Then, the spectral Christoffel–Geronimus–Uvarov formulas of Theorem 6.17 can be applied putting $\mathcal{T} = 0$.

6.6.3 More transformations

We could consider a slightly more general situation with transformations of the following non-symmetric form, where W has a nonsingular leading coefficient,

$$\hat{u}=W(x)uV(x)=W_C(x)u(W_G(x))^{-1},\qquad W_C(x)=\det(V(x))W(x),\qquad W_G(x)=\operatorname{adj}(V(x)),$$

or

$$\hat{u}=(W(x))^{-1}u(V(x))^{-1}=W_C(x)u(W_G(x))^{-1},\quad W_C(x)=\operatorname{adj}(W(x)),\quad W_G(x)=\det(W(x))V(x).$$

In this case the polynomial $V(x)$ can have a singular leading coefficient, and, in such a situation we apply the mixed spectral/nonspectral Christoffel–Geronimus–Uvarov formulas.

6.6.4 Degree one matrix Geronimus-Uvarov transformations

We consider perturbing polynomials of degree one, i.e.

$$W_C(x)=xI_p-A_C,\qquad\qquad W_G(x)=xI_p-A_G,$$

and massless situation, i.e. all $\xi^{[a]}_{j,m}=0$ in (6.1). For the Jordan pairs (X_C,J_C) and (X_G,J_G) we have $A_C=X_CJ_C(X_C)^{-1}$ and $A_G=X_GJ_G(X_G)^{-1}$. Now, for $n\geq N_G$, from Theorem 6.17, the following last quasi-determinant expressions hold

$$\begin{aligned}
\hat{P}^{[1]}_n(x)(xI_p-A_C)&=\Theta_*\begin{bmatrix}P^{[1]}_{n-1}(A_C)X_C & C^{[1]}_{n-1}(A_G)X_G & P^{[1]}_{n-1}(x)\\ P^{[1]}_{n}(A_C)X_C & C^{[1]}_{n}(A_G)X_G & P^{[1]}_{n}(x)\\ P^{[1]}_{n+1}(A_C)X_C & C^{[1]}_{n+1}(A_G)X_G & P^{[1]}_{n+1}(x)\end{bmatrix},\\
\hat{H}_n&=\Theta_*\begin{bmatrix}P^{[1]}_{n-1}(A_C)X_C & C^{[1]}_{n-1}(A_G)X_G & H_{n-1}\\ P^{[1]}_{n}(A_C)X_C & C^{[1]}_{n}(A_G)X_G & 0_p\\ P^{[1]}_{n+1}(A_C)X_C & C^{[1]}_{n+1}(A_G)X_G & 0_p\end{bmatrix},\\
\big(\hat{P}^{[2]}_n(y)\big)^\dagger&=-\Theta_*\begin{bmatrix}P^{[1]}_{n-1}(A_C)X_C & C^{[1]}_{n-1}(A_G)X_G & H_{n-1}\\ P^{[1]}_{n}(A_C)X_C & C^{[1]}_{n}(A_G)X_G & 0_p\\ (yI_p-A_G)K_{n-1}(A_C,y)X_C & \big(yI_p-A_G\big)K^{(pc)}_{n-1}(A_G,y)+I_p\big)X_G & 0_p\end{bmatrix}.
\end{aligned}$$

If we expand the quasi-determinant we get

$$\begin{aligned}
\hat{P}^{[1]}_n(x)(xI_p-A_C)=&P^{[1]}_{n+1}(x)-C^{[1]}_{n+1}(A_G)\Big(C^{[1]}_{n}(A_G)-P^{[1]}_{n}(A_C)\big(P^{[1]}_{n-1}(A_C)\big)^{-1}C^{[1]}_{n-1}(A_G)\Big)^{-1}P^{[1]}_{n}(x)\\
&-C^{[1]}_{n+1}(A_G)\Big(C^{[1]}_{n-1}(A_G)-P^{[1]}_{n-1}(A_C)\big(P^{[1]}_{n}(A_C)\big)^{-1}C^{[1]}_{n}(A_G)\Big)^{-1}P^{[1]}_{n-1}(x)\\
&-P^{[1]}_{n+1}(A_C)\Big(P^{[1]}_{n}(A_C)-C^{[1]}_{n}(A_G)\big(C^{[1]}_{n-1}(A_G)\big)^{-1}P^{[1]}_{n-1}(A_C)\Big)^{-1}P^{[1]}_{n}(x)\\
&-P^{[1]}_{n+1}(A_C)\Big(P^{[1]}_{n-1}(A_C)-C^{[1]}_{n-1}(A_G)\big(C^{[1]}_{n}(A_G)\big)^{-1}P^{[1]}_{n}(A_C)\Big)^{-1}P^{[1]}_{n-1}(x)
\end{aligned}$$

$$\begin{aligned}\hat{H}_n =&-\Big(C^{[1]}_{n+1}(A_G)\Big(C^{[1]}_{n-1}(A_G)-P^{[1]}_{n-1}(A_C)\big(P^{[1]}_n(A_C)\big)^{-1}C^{[1]}_n(A_G)\Big)^{-1}\\&+P^{[1]}_{n+1}(A_C)\Big(P^{[1]}_{n-1}(A_C)-C^{[1]}_{n-1}(A_G)\big(C^{[1]}_n(A_G)\big)^{-1}P^{[1]}_n(A_C)\Big)^{-1}\Big)H_{n-1},\\\big(\hat{P}^{[2]}_n(y)\big)^\dagger =&\Big(\big((yI_p-A_G)K^{(pc)}_{n-1}(A_G,y)+I_p\big)\Big(C^{[1]}_{n-1}(A_G)-P^{[1]}_{n-1}(A_C)\big(P^{[1]}_n(A_C)\big)^{-1}C^{[1]}_n(A_G)\Big)^{-1}\\&+(yI_p-A_G)K_{n-1}(A_C,y)\Big(P^{[1]}_{n-1}(A_C)-C^{[1]}_{n-1}(A_G)\big(C^{[1]}_n(A_G)\big)^{-1}P^{[1]}_n(A_C)\Big)^{-1}\Big)H_{n-1}.\end{aligned}$$

6.7 Matrix Uvarov transformations with a finite discrete support

Uvarov perturbations in the scalar context with a finite number of Dirac deltas, i.e. linear functionals supported at different points, have been considered first in [143]. For the matrix case, it was studied in a series of papers [145, 146, 147] where the corresponding Christoffel–Geronimus–Uvarov formula for the perturbed polynomials, when a solely Dirac delta is added in a point, are deduced. In this Section we present the general case when we have an additive perturbation of discrete finite support, allowing therefore for an arbitrary finite number of derivatives of the Dirac delta at several different points. We have two reasons to include this material in this Section. Our main motivations are, first to show how some of the tools, like spectral jets, used in the body of this dissertation also apply in this context and, second, to achieve a more complete account of the family of transformations of Darboux type for matrix orthogonal polynomials. For the multivariate scenario these transformations have been discussed in [21, 43, 44].

Proposition 6.37 (Additive perturbation and reproducing kernels)**.** *Let us consider an additive perturbation $\hat{u}$ of the matrix of linear functional u defined by*

$$\hat{u}=u+v,$$

and let us assume that u and $\hat{u}$ are quasi-definite. Then,

$$\hat{P}^{[1]}_n(x)=P^{[1]}_n(x)-\big\langle\hat{P}^{[1]}_n(y),\big(K_{n-1}(x,y)\big)^\dagger\big\rangle_v,\quad \big(\hat{P}^{[2]}_n(x)\big)^\dagger=\big(P^{[2]}_n(x)\big)^\dagger-\big\langle K_{n-1}(y,x),\hat{P}^{[2]}_n(y)\big\rangle_v,$$

and

$$\hat{H}_n=H_n+\Big\langle\hat{P}^{[1]}_n(y),P^{[2]}_n(y)\Big\rangle_v.$$

Proof. From (1.11) and (1.12)

$$\Big\langle\hat{P}^{[1]}_n(y),P^{[2]}_m(y)\Big\rangle_{\hat{u}}=0_p,\quad\Big\langle P^{[1]}_m(y),\hat{P}^{[2]}_n(y)\Big\rangle_{\hat{u}}=0_p,\quad m\in\{1,\dots n-1\},\tag{6.28}$$

$$\Big\langle\hat{P}^{[1]}_n(y),P^{[2]}_n(y)\Big\rangle_{\hat{u}}=\hat{H}_n=\Big\langle P^{[1]}_n(y),\hat{P}^{[2]}_n(y)\Big\rangle_{\hat{u}}.\tag{6.29}$$

For $m \in \{1,\dots n-1\}$, (6.28) can be expressed as

$$\left\langle \hat{P}^{[1]}_n(y), P^{[2]}_m(y)\right\rangle_u = -\left\langle \hat{P}^{[1]}_n(y), P^{[2]}_m(y)\right\rangle_v, \qquad \left\langle P^{[1]}_m(y), \hat{P}^{[2]}_n(y)\right\rangle_u = -\left\langle P^{[1]}_m(y), \hat{P}^{[2]}_n(y)\right\rangle_v.$$

Then, recalling (1.13) we get

$$\begin{aligned}\left\langle \hat{P}^{[1]}_n(y), (K_{n-1}(x,y))^{\dagger}\right\rangle_u &= -\left\langle \hat{P}^{[1]}_n(y), (K_{n-1}(x,y))^{\dagger}\right\rangle_v,\\ \left\langle K_{n-1}(y,x), \hat{P}^{[2]}_n(y)\right\rangle_u &= -\left\langle K_{n-1}(y,x), \hat{P}^{[2]}_n(y)\right\rangle_v.\end{aligned}$$

But $\hat{P}^{[1]}_n(y) - P^{[1]}_n(y)$ and $\hat{P}^{[2]}_n(y) - P^{[2]}_n(y)$ have degree $n-1$. Therefore, recalling (1.14) we deduce

$$\begin{aligned}\hat{P}^{[1]}_n(x) - P^{[1]}_n(x) &= \left\langle \hat{P}^{[1]}_n(y) - P^{[1]}_n(y), (K_{n-1}(x,y))^{\dagger}\right\rangle_u = \left\langle \hat{P}^{[1]}_n(y), (K_{n-1}(x,y))^{\dagger}\right\rangle_u\\ &= -\left\langle \hat{P}^{[1]}_n(y), (K_{n-1}(x,y))^{\dagger}\right\rangle_v,\\ \left(\hat{P}^{[2]}_n(x) - P^{[2]}_n(x)\right)^{\dagger} &= \left\langle K_{n-1}(y,x), \hat{P}^{[2]}_n(y) - P^{[2]}_n(y)\right\rangle_u = \left\langle K_{n-1}(y,x), \hat{P}^{[2]}_n(y)\right\rangle_u\\ &= -\left\langle K_{n-1}(y,x), \hat{P}^{[2]}_n(y)\right\rangle_v.\end{aligned}$$

Finally, form (6.29) we deduce

$$\begin{aligned}\hat{H}_n &= \left\langle \hat{P}^{[1]}_n(y), P^{[2]}_n(y)\right\rangle_{\hat{u}} = \left\langle \hat{P}^{[1]}_n(y), P^{[2]}_n(y)\right\rangle_u + \left\langle \hat{P}^{[1]}_n(y), P^{[2]}_n(y)\right\rangle_v\\ &= H_n + \left\langle \hat{P}^{[1]}_n(y), P^{[2]}_n(y)\right\rangle_v.\end{aligned}$$

■

6.7.1 Finite discrete support additive perturbations

We now assume that the additive perturbation has a finite support. Thus, see [74, 131],

$$v = \sum_{a=1}^{q}\sum_{m=0}^{\kappa^{(a)}-1} \frac{(-1)^m}{m!}\delta^{(m)}(x-x_a)\beta^{(a)}_m, \qquad \beta^{(a)}_m \in \mathbb{C}^{p\times p}.$$

We will assume along this subsection that both u and $u+v$ are quasi-definite matrices of linear functionals.

Definition 6.38. *For the family of perturbation matrices $\beta^{(a)}_m \in \mathbb{C}^{p\times p}$ we introduce*

$$\beta^{(a)} := \begin{bmatrix} \beta^{(a)}_0 & \beta^{(a)}_1 & \beta^{(a)}_2 & & & \beta^{(a)}_{\kappa^{(a)}-1}\\ \beta^{(a)}_1 & \beta^{(a)}_2 & \iddots & & \beta^{(a)}_{\kappa^{(a)}-1} & 0_p\\ \beta^{(a)}_2 & \iddots & & \beta^{(a)}_{\kappa^{(a)}-1} & 0_p & 0_p\\ \iddots & & \iddots & & & \\ \beta^{(a)}_{\kappa^{(a)}-1} & 0_p & 0_p & 0_p & 0_p & 0_p \end{bmatrix} \in \mathbb{C}^{\kappa^{(a)}p\times\kappa^{(a)}p}$$

and if $N := \kappa^{(1)} + \cdots + \kappa^{(q)}$, *let*

$$\beta = \operatorname{diag}(\beta^{(1)}, \dots, \beta^{(q)}) \in \mathbb{C}^{Np \times Np}.$$

Definition 6.39. *The spectrum of the perturbation is its support taking into account the order of each linear functional involved, that is*

$$\sigma(v) = \{(x_a, \kappa^{(a)})\}_{a=1}^{q}.$$

The spectral jet associated with the finite support matrix of linear functionals v is, for any sufficiently smooth matrix function $f(x)$ *defined in an open set in* $\mathbb{C}$,

$$\mathsf{J}_f = \left[f(x_1), \dots, \frac{(f(x))_{x_1}^{(\kappa^{(1)}-1)}}{(\kappa^{(1)}-1)!}, \dots, f(x_q), \dots, \frac{(f(x))_{x_q}^{(\kappa^{(q)}-1)}}{(\kappa^{(q)}-1)!} \right] \in \mathbb{C}^{p \times Np}.$$

For bivariate matrix functions $K(x,y)$ *we have three different types of spectral jets, the partial spectral jet with respect to the first variable*

$$\mathsf{J}_K^{[1,0]}(y) = \left[K(x_1,y), \dots, \frac{(K(x,y))_{x_1,y}^{(\kappa^{(1)}-1,0)}}{(\kappa^{(1)}-1)!}, \dots, K(x_q,y), \dots, \frac{(K(x,y))_{x_q,y}^{(\kappa^{(q)}-1,0)}}{(\kappa^{(q)}-1)!} \right] \in \mathbb{C}^{p \times Np},$$

and also with respect to the second variable

$$\mathsf{J}_K^{[0,1]}(x) = \begin{bmatrix} K(x,x_1) \\ \vdots \\ \dfrac{(K(x,y))_{x,x_1}^{(0,\kappa^{(1)}-1)}}{(\kappa^{(1)}-1)!} \\ \vdots \\ K(x,x_q) \\ \vdots \\ \dfrac{(K(x,y))_{x,x_q}^{(0,\kappa^{(q)}-1)}}{(\kappa^{(q)}-1)!} \end{bmatrix} \in \mathbb{C}^{Np \times p}.$$

We also introduce a spectral jet with respect to both two variables as follows

$$\mathtt{J}_K=\begin{bmatrix} K(x_1,x_1) & \dots & \frac{(K(x,y))^{(\kappa^{(1)}-1,0)}_{x_1,x_1}}{(\kappa^{(1)}-1)!} & \dots & K(x_q,x_1) & \dots & \frac{(K(x,y))^{(\kappa^{(q)}-1,0)}_{x_q,x_1}}{(\kappa^{(q)}-1)!}\\ \vdots & & \vdots & & \vdots & & \vdots\\ \frac{(K(x,y))^{(0,\kappa^{(1)}-1)}_{x_1,x_1}}{(\kappa^{(1)}-1)!} & \dots & \frac{(K(x,y))^{(\kappa^{(1)}-1,\kappa^{(1)}-1)}_{x_1,x_1}}{(\kappa^{(1)}-1)!(\kappa^{(1)}-1)!} & \dots & \frac{(K(x,y))^{(0,\kappa^{(1)}-1)}_{x_q,x_1}}{(\kappa^{(1)}-1)!} & \dots & \frac{(K(x,y))^{(\kappa^{(q)}-1,\kappa^{(1)}-1)}_{x_1,x_q}}{(\kappa^{(q)}-1)!(\kappa^{(1)}-1)!}\\ \vdots & & \vdots & & \vdots & & \vdots\\ K(x_1,x_q) & \dots & \frac{(K(x,y))^{(\kappa^{(1)}-1,0)}_{x_1,x_q}}{(\kappa^{(1)}-1)!} & \dots & K(x_q,x_q) & \dots & \frac{(K(x,y))^{(\kappa^{(q)}-1,0)}_{x_q,x_q}}{(\kappa^{(q)}-1)!}\\ \vdots & & \vdots & & \vdots & & \vdots\\ \frac{(K(x,y))^{(0,\kappa^{(q)}-1)}_{x_1,x_q}}{(\kappa^{(q)}-1)!} & \dots & \frac{(K(x,y))^{(\kappa^{(1)}-1,\kappa^{(q)}-1)}_{x_1,x_q}}{(\kappa^{(1)}-1)!(\kappa^{(q)}-1)!} & \dots & \frac{(K(x,y))^{(0,\kappa^{(q)}-1)}_{x_q,x_q}}{(\kappa^{(q)}-1)!} & \dots & \frac{(K(x,y))^{(\kappa^{(q)}-1,\kappa^{(q)}-1)}_{x_q,x_q}}{(\kappa^{(q)}-1)!(\kappa^{(q)}-1)!} \end{bmatrix},$$

which belongs to $\mathbb{C}^{Np\times Np}$*. We have used the compact notation*

$$(K(x,y))^{(n,m)}_{a,b}=\frac{\partial^{n+m}K}{\partial x^n\partial y^m}\Big|_{x=a,y=b}.$$

Proposition 6.40. *The following relations hold*

$$\begin{aligned}\left\langle \hat{P}^{[1]}_n(y),(K_{n-1}(x,y))^{\dagger}\right\rangle_v&=\mathtt{J}_{\hat{P}^{[1]}_n}\beta\mathtt{J}^{[0,1]}_{K_{n-1}}(x),\\ \left\langle K_{n-1}(y,x),\hat{P}^{[2]}_n(y)\right\rangle_v&=\mathtt{J}^{[1,0]}_{K_{n-1}}(x)\beta\big(\mathtt{J}_{\hat{P}^{[2]}_n}\big)^{\dagger}.\end{aligned}$$

Proof. We prove only the first one, since the second one is deduced in a similar way. A direct substitution gives

$$\left\langle \hat{P}^{[1]}_n(y),(K_{n-1}(x,y))^{\dagger}\right\rangle_v=\sum_{a=1}^{q}\sum_{m=0}^{\kappa^{(a)}-1}\frac{1}{m!}\Big(\hat{P}^{[1]}_n(y)\beta^{(a)}_m K_{n-1}(x,y)\Big)^{(m)}_{y=x_a},$$

but

$$\frac{1}{m!}\Big(\hat{P}^{[1]}_n(y)\beta^{(a)}_m K_{n-1}(x,y)\Big)^{(m)}_{y=x_a}=\sum_{k=0}^{m}\frac{1}{(m-k)!}\big(\hat{P}^{[1]}_n(y)\big)^{(m-k)}_{y=x_a}\beta^{(a)}_m\frac{1}{k!}\big(K_{n-1}(x,y)\big)^{(k)}_{y=x_a},$$

and the result follows. ■

Theorem 6.41. *The perturbed matrix orthogonal polynomials and quasitau matrices are*

$$\hat{P}^{[1]}_n(x)=\Theta_*\begin{bmatrix} I_{Np}+\beta\mathtt{J}_K & \beta\mathtt{J}^{[0,1]}_{K_{n-1}}(x)\\ \mathtt{J}_{P^{[1]}_n} & P^{[1]}_n(x)\end{bmatrix},$$

$$(\hat{P}_n^{[2]}(x))^\dagger = \Theta_* \begin{bmatrix} I_{Np} + \beta \mathtt{J}_K & \beta(\mathtt{J}_{P_n^{[2]}})^\dagger \\ \mathtt{J}_{K_{n-1}}^{[1,0]}(x) & (P_n^{[2]}(x))^\dagger \end{bmatrix},$$
$$\hat{H}_n = \Theta_* \begin{bmatrix} I_{Np} + \beta \mathtt{J}_K & -\beta(\mathtt{J}_{P_n^{[2]}})^\dagger \\ \mathtt{J}_{P_n^{[1]}} & H_n \end{bmatrix}.$$

Proof. Now, the comparison of Propositions 6.37 and 6.40 yields

$$\begin{aligned} \hat{P}_n^{[1]}(x) &= P_n^{[1]}(x) - \mathtt{J}_{\hat{P}_n^{[1]}} \beta \mathtt{J}_{K_{n-1}}^{[0,1]}(x), \\ (\hat{P}_n^{[2]}(x))^\dagger &= (P_n^{[2]}(x))^\dagger - \mathtt{J}_{K_{n-1}}^{[1,0]}(x) \beta \big(\mathtt{J}_{\hat{P}_n^{[2]}}\big)^\dagger. \end{aligned} \tag{6.30}$$

Again, we take spectral jets to get

$$\begin{aligned} \mathtt{J}_{\hat{P}_n^{[1]}} &= \mathtt{J}_{P_n^{[1]}} - \mathtt{J}_{\hat{P}_n^{[1]}} \beta \mathtt{J}_{K_{n-1}} \\ (\mathtt{J}_{\hat{P}_n^{[2]}})^\dagger &= (\mathtt{J}_{P_n^{[2]}})^\dagger - \mathtt{J}_{K_{n-1}} \beta \big(\mathtt{J}_{\hat{P}_n^{[2]}}\big)^\dagger, \end{aligned}$$

and, consequently,

$$\begin{aligned} \mathtt{J}_{\hat{P}_n^{[1]}}(I_{Np} + \beta \mathtt{J}_{K_{n-1}}) &= \mathtt{J}_{P_n^{[1]}} \\ \beta^{-1}(I_{Np} + \beta \mathtt{J}_{K_{n-1}}) \beta (\mathtt{J}_{\hat{P}_n^{[2]}})^\dagger &= (\mathtt{J}_{P_n^{[2]}})^\dagger. \end{aligned} \tag{6.31}$$

Let us check that the matrix $I_{Np} + \beta \mathcal{J}_{K_{n-1}}$ is nonsingular from the quasi-definiteness of u. Indeed, if we assume the contrary, we can find a nonzero vector $X \in \mathbb{C}^{Np}$ such that $(I_{Np} + \beta \mathcal{J}_{K_{n-1}})X = 0$, the zero vector in $\mathbb{C}^{Np}$. Thus, using (6.31) we conclude that $\mathtt{J}_{P_n^{[1]}} X = 0$. But, taking into account

$$I_{Np} + \beta \mathcal{J}_{K_n} = I_{Np} + \beta \mathcal{J}_{K_{n-1}} + \beta (\mathtt{J}_{P_n^{[2]}})^\dagger (H_n)^{-1} \mathtt{J}_{P_n^{[1]}},$$

by induction, we deduce that $\mathtt{J}_{P_k^{[1]}} X = 0$ for $k \in \{n, n+1, \dots\}$. If

$$X = \begin{bmatrix} X_0^{(1)} \\ \vdots \\ X_{\kappa^{(1)}-1}^{(1)} \\ \vdots \\ X_0^{(q)} \\ \vdots \\ X_{\kappa^{(q)}-1}^{(q)} \end{bmatrix},$$

then the linear functional

$$v_X = \sum_{a=1}^{q} \sum_{m=0}^{\kappa^{(a)}-1} \frac{(-1)^m}{m!} \delta^{(m)}(x - x_a) X_m^{(a)}$$

satisfies (see Definition 1.48)

$$\langle P_k^{[1]}(x), v_X \rangle = 0$$

for $k \in \{n, n+1, \dots\}$. Then

$$v_X \in \left((P_k^{[1]}(x))_{k\in\mathbb{N}}\right)^{\perp} := \left\{\tilde{u} \text{ a matrix of linear functionals such that } \langle P_k^{[1]}(x), \tilde{u}\rangle = 0_p : k \geq n\right\}.$$

At this point it is convenient to recall that the topological and algebraic duals of the set of matrix polynomials coincide i.e $(\mathbb{C}^{p\times p}[x])' = (\mathbb{C}^{p\times p}[x])^* = \mathbb{C}^{p\times p}[\![x]\!]$, where we understand the set of matrix polynomials or matrix formal series as left or right modules over the ring of matrices. We also recall that the module of matrix polynomials of degree less than or equal to m is isomorphic to the free module $\left(\mathbb{C}^{p\times p}\right)^m$, a Cartesian product of m copies of the ring of $p\times p$ complex matrices. Therefore, for each positive integer m, we consider the linear basis given by the following set of matrices of linear functionals $((P_k^{[1]})^*)_{k=0}^m$, the dual of $(P_k^{[1]}(x))_{k=0}^m$, given by $\langle P_k^{[1]}(x), (P_l^{[1]})^*\rangle = \delta_{k,l} I_p$, such that any matrix of linear functionals can be written $\sum_{k=0}^m (P_l^{[1]})^* C_k$, where $C_k \in \mathbb{C}^{p\times p}$. Then, $\left((P_k^{[1]}(x))_{k\in\mathbb{N}}\right)^{\perp} = \left((P_k^{[1]})^*\right)_{k=0}^{n-1}\mathbb{C}^{p\times p} \cong \left(\mathbb{C}^{p\times p}\right)^n$, and, therefore, it is a free right module. But, according to (1.10), the set of matrices of linear functionals $\left(u\left(P_k^{[2]}(x)\right)^{\dagger}\right)_{k=0}^{n-1} \subseteq \left(\left(P_k^{[1]}(x)\right)_{k=n}^{\infty}\right)^{\perp} = \left((P_k^{[1]})^*\right)_{k=0}^{n-1}\mathbb{C}^{p\times p}$, so that

$$\left(u\left(P_k^{[2]}(x)\right)^{\dagger}\right)_{k=0}^{n-1}\mathbb{C}^{p\times p} = \left(\left(P_k^{[1]}(x)\right)_{k=n}^{\infty}\right)^{\perp}.$$

Thus, we can write $v_X = uQ(x)$, where $Q(x)$ is a matrix polynomial and $\deg Q(x) \leq n-1$. For the matrix polynomial $R(x) := (x-x_1)^{\kappa^{(1)}} \cdots (x-x_q)^{\kappa^{(q)}} I_p$, we get $v_X R(x) = 0$. This implies that

$$uQ(x)R(x) = 0.$$

Hence, u is not quasi-definite, as the polynomial $Q(x)R(x)$ is orthogonal to the set of matrix polynomials. This fact contradicts our hypothesis, and, consequently, the matrix $I_{Np} + \beta \mathtt{J}_{K_{n-1}}$ is nonsingular, allowing to write (6.31) as

$$\mathtt{J}_{\hat{P}_n^{[1]}} = \mathtt{J}_{P_n^{[1]}} (I_{Np} + \beta \mathtt{J}_{K_{n-1}})^{-1}.$$

This relation, when introduced in (6.30), gives

$$\hat{P}_n^{[1]}(x) = P_n^{[1]}(x) - \mathtt{J}_{P_n^{[1]}} (I_{Np} + \beta \mathtt{J} J_{K_{n-1}})^{-1} \beta \mathtt{J}_{K_{n-1}}^{[0,1]}(x),$$

and the result follows. Finally, from Proposition 6.37 we get

$$\begin{aligned}\hat{H}_n &= H_n + \mathtt{J}_{\hat{P}_n^{[1]}} \beta \left(\mathtt{J}_{P^{[2]}}\right)^{\dagger} \\ &= H_n + \mathtt{J}_{P_n^{[1]}} (I_{Np} + \beta \mathtt{J}_{K_{n-1}})^{-1} \beta \left(\mathtt{J}_{P^{[2]}}\right)^{\dagger}.\end{aligned}$$

■

Chapter 7

Conclusions and future research

7.1 Main contributions

Here we summarize the main contributions of this Ph.D. thesis

- We consider the multiple Geronimus transformation and show that it yields a discrete (non-diagonal) Sobolev inner product. Moreover, we show that every Sobolev inner product can be obtained as a multiple Geronimus transformation of a measure. Based on the Durán's papers, we also find a close relation between the multiple and matrix Geronimus transformations.

- The above results give us a motivation to study the matrix (symmetric) Geronimus transformation. Here we find conditions for the existence of the sequence of monic matrix orthogonal polynomials (perturbed) with respect to the new sesquilinear form associated with the matrix Geronimus transformation, as well as a connection formula between the sequences of perturbed and the original monic matrix orthogonal polynomials.

- From a more general framework, we study three matrix transformations of a matrix of linear functionals, u. They are the Christoffel matrix transformation ($\hat{u} = W_C(x)u$), the Geronimus matrix transformation ($W_G(x)\check{u} = u$), and the Geronimus-Uvarov matrix transformation ($W_G(x)\hat{u} = uW_C(x)$), where $W_C(x)$ and $W_G(x)$ are two matrix polynomials. For the Christoffel matrix transformation we give the connection formula between the original and perturbed matrix bi- orthogonal polynomials when the leading coefficient of $W_C(x)$ is a nonsingular matrix.

  For the Geronimus and Geronimus-Uvarov matrix transformation, we obtain a representation of their corresponding sequences of matrix bi-orthogonal polynomials using spectral and non spectral methods. In the spectral method, we find the representation of the perturbed bi-orthogonal polynomials depending on the family of original ones and the second kind functions. Here, we use the fact that the leading coefficients of $W_C(x)$ and $W_G(x)$ are nonsingular matrices. In the non spectral method, we give a representation of the perturbed bi-orthogonal polynomials without any assumption concerning the leading coefficient of $W_G(x)$. These representations are only given in terms of the original bi-orthogonal polynomials. However, we pay the penalty of the application of the perturbed linear functional on the original bi-orthogonal polynomials multiplies by a monomial.

- As applications, we study the extension of Christoffel matrix transformations to non-Abelian 2D Toda hierarchies. Besides, we study the matrix Uvarov transformations, which are a special case of Geronimus-Uvarov transformations.

The original results of this thesis have been published in several international research Journals

[11] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas, F. Marcellán, *Christoffel transformations for matrix orthogonal polynomials in the real line and the non-Abelian 2D Toda lattice hierarchy*, Internat. Math. Res. Notices, **5**, (2017) 1285-1341.

[12] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas, F. Marcellán, *Transformation theory and Christoffel formulas for matrix biorthogonal polynomials on the real line*, arXiv:1605.04617v7 [math.CA].

[47] M. Derevyagin, J. C. García-Ardila, F. Marcellán, *Multiple Geronimus transformations.* Linear Algebra Appl. **454** (2014), 158-183.

[67] J. C. García-Ardila, L. E. Garza, F. Marcellán, *A canonical Geronimus transformation for matrix orthogonal polynomials,* Linear Multilinear Algebra doi:10.1080/03081087.2017.1299089.

[68] J. C. García-Ardila, L. E. Garza, F. Marcellán, *An Extension of the Geronimus Transformation for Orthogonal Matrix Polynomials on the Real Line* Mediterr. J. Math. **13** (2016), 5009-5032.

7.2 Open problems

Finally, we discuss some related work as well a set of open problems for a future research.

P.1 In Subsection 6.6.2 we deal with a symmetric linear functional u such that we apply a symmetric Christoffel transformation i.e. $\hat{u} = W_C(x)uW_C^\top(x)$, where $W_C(x)$ is a matrix polynomial. In order to obtain a representation of the matrix orthogonal polynomials associated with $\hat{u}$ we use the fact that the leading coefficient of $W_C(x)$ is a nonsingular matrix. But what happens if this hypothesis fails?. A similar problem can be formulated for symmetric Geronimus transformations.

P.2 In this thesis we study three families of transformations. However we were not exhaustive. Indeed, what can be said when for a matrix of linear functionals we first apply a matrix Geronimus (Christoffel) transformation to left (right) and then a Christoffel (Geronimus) transformation also to left (right)?.

P.3 Give a right definition of linear spectral transformation of matrix linear functionals. Describe generator set of such linear transformations.

P.4 The theory of spectral matrix transformations analyzed in this thesis can be extended to positive definite matrices of measures supported on the unit circle following the program for the scalar case described in [29].

Bibliography

[1] M. Adler, P. van Moerbeke, *Group factorization, moment matrices and Toda lattices*, Internat. Math. Res. Notices, **12** (1997), 556-572.

[2] M. Adler, P. van Moerbeke, *Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems*, Commun. Math. Phys. **207** (1999), 589-620.

[3] M. Adler, P. van Moerbeke, *Darboux transforms on band matrices, weights and associated polynomials*, Internat. Math. Res. Notices **18** (2001), 935-984.

[4] M. Adler, P. van Moerbeke, P. Vanhaecke, *Moment matrices and multi- component KP, with applications to random matrix theory*, Commun. Math. Phys. **286** (2009), 1-38.

[5] N. I. Akhiezer, *Classical moment problem and related questions in analysis,* Oliver & Boyd, Edinburgh, 1965.

[6] M. Alfaro, F. Marcellán, M. L. Rezola, *Orthogonal polynomials on Sobolev spaces: Old and new directions,* J. Comput. Appl. Math. **48** (1993), 113-131.

[7] M. Alfaro, F. Marcellán, A. Peña, M. L. Rezola, *Orthogonal polynomials associated with an inverse quadratic spectral transform*, Comput. Math. Appl. **61** (2011), 888-900.

[8] M. Alfaro, F. Marcellán, M. L. Rezola, A. Ronveaux, *On orthogonal polynomials of Sobolev type: Algebraic properties and zeros*, SIAM J. Math. Anal. **23** (1992), 737-757.

[9] M. Alfaro, F. Marcellán, M. L. Rezola, A. Ronveaux, *Sobolev-Type orthogonal polynomials: The nondiagonal case*, J. Approx. Theory **83** (1995), 266-287.

[10] M. Alfaro, F. Marcellán, A. Peña, M. L. Rezola, *When do linear combinations of orthogonal polynomials yield new sequences of orthogonal polynomials?*, J. Comput. Appl. Math. **233** (2010), 1446-1452.

[11] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas, F. Marcellán, *Christoffel transformations for matrix orthogonal polynomials in the real line and the non-Abelian 2D Toda lattice hierarchy*, Internat. Math. Res. Notices, **5** (2017), 1285-1341.

[12] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas, F. Marcellán, *Transformation theory and Christoffel formulas for matrix biorthogonal polynomials on the real line*, arXiv:1605.04617v7 [math.CA].

[13] C. Álvarez-Fernández, M. Mañas, Fidalgo Prieto, *The multicomponent 2D Toda hierarchy: generalized matrix orthogonal polynomials, multiple orthogonal polynomials and Riemann-Hilbert problems*, Inverse Probl. **26** (2010), 055009 (15pp).

[14] C. Álvarez-Fernández, U. Fidalgo Prieto, M. Mañas, *Multiple orthogonal polynomials of mixed type: Gauss-Borel factorization and the multi-component 2D Toda hierarchy* , Adv. Math. **227** (2011), 1451-1525.

[15] C. Álvarez-Fernández, M. Mañas, *On the Christoffel-Darboux formula for generalized matrix orthogonal polynomials*, J. Math. Anal. Appl. **418** (2014), 238-247.

[16] R. Álvarez-Nodarse, A. J. Durán, A. M. de los Ríos, *Orthogonal matrix polynomials satisfying second order difference equation,* J. Approx. Theory **169** (2013), 40-55.

[17] G. E. Andrews, R. Askey, R. Roy, *Special Functions*, Encyclopedia of Mathematics and its applications **71**, Cambridge University Press, Cambridge, 1999.

[18] A.I. Aptekarev, E. M. Nikishin, *The scattering problem for a discrete Sturm-Liouville operator*, Mat. Sb. **121** (163): 327-358 (1983); Math. USSR Sb. **49** (1984), 325-355.

[19] G. Ariznabarreta, M. Mañas, *Matrix orthogonal Laurent polynomials on the unit circle and Toda type integrable systems,* Adv. Math. **264** (2014), 396-463.

[20] G. Ariznabarreta, M. Mañas, *Multivariate orthogonal polynomials and integrable systems*, Adv. Math. **302** (2016), 628-739.

[21] G. Ariznabarreta, M. Mañas, *Linear spectral transformations for multivariate orthogonal polynomials and multispectral Toda hierarchies* arXiv:1511.09129.

[22] D. Beghdadi, P. Maroni, *On the inverse problem of the product of a semi-classical form by a polynomial*, J. Comput. Appl. Math. **88** (1998), 377-399.

[23] M. J. Bergvelt, A. P. E. ten Kroode, *Tau-functions and zero-curvature equations of Toda-AKNS type*, J. Math. Phys. **29** (1988), 1308-1320.

[24] M. J. Bergvelt, A. P. E. ten Kroode, *Partitions, Vertex Operators Constructions and Multi-Component KP Equations*, Pacific. J. Math. **171** (1995), 23-88.

[25] R. P. Boas, *The Stieltjes moment problem for functions of bounded variation* Bull. Amer. Math. Soc. **45** (1939), 399-404.

[26] A. Branquinho, F. Marcellán, *Generating new classes of orthogonal polynomials*, Internat. J. Math. Math. Sci. **19** (1996), 643-656.

[27] M. I. Bueno, F. Marcellán, *Darboux transformation and perturbation of linear functionals*, Linear Algebra Appl. **384** (2004), 215-242.

[28] M. I. Bueno, F. Marcellán, *Polynomial perturbations of bilinear functionals and Hessenberg matrices*, Linear Algebra Appl. **414** (2006), 64-83.

[29] M. J. Cantero, F Marcellán, L. Moral, L. Velázquez *Darboux transfromations for CMV matrices,* Adv. Math. **298** (2016), 122-206.

[30] A. D. D. Craik, *Prehistory of Faà di Bruno's formula,* Amer. Math. Monthly **112** (2005), no. 2, 119-130.

[31] M. M. Castro, F. A. Grünbaum, *Orthogonal matrix polynomials satisfying first order differential equations: a collection of instructive examples,* J. Nonlinear Math. Phys. **12** (2005), suppl. 2, 63-76.

[32] V. S. Chelyshkov, *Alternative Jacobi polynomials and orthogonal exponentials*, arXiv:1105.1838.

[33] T. S. Chihara, *An Introduction to Orthogonal Polynomials*. In : *Mathematics and its Applications Series*, Vol. **13**. Gordon and Breach Science Publishers, New York-London-Paris, 1978.

[34] T. S. Chihara, *Orthogonal Polynomials and measures with end point masses*, Rocky Mountain J. Math. **15** (1985), 705-719.

[35] A. E. Choque Rivero, L. E. Garza, *Moment perturbation of matrix polynomials*, Integral Transforms Spec. Funct. **26** (2015), 177-191.

[36] E. B. Christoffel, *Über die Gaussische Quadratur und eine Verallgemeinerung derselben*, J. Reine Angew. Math. **55** (1858), 61-82.

[37] R. G. Cooke, *Infinite matrices and sequence spaces,* Dover Publications, Inc., New York 1965.

[38] D.Damanik, A. Pushnitski, B. Simon, *The analytic theory of matrix orthogonal polynomials*, Surveys in Approx. Theory **4** (2008), 1-85.

[39] G. Darboux, *Mémoire sur l' approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série*, J. Math. Pur. Appl. (3) **4** (1878), 5-56; 377-416.

[40] G. Darboux, *Sur une proposition relative aux équations linéaires*, Comptes Rendus hebd. séances Acad. Sci. Paris **94** (1882), 1456-1459.

[41] M. De la Iglesia, A. Durán. *Some examples of orthogonal matrix polynomials satisfying odd order differential equations*, J. Approx. Theory **150** (2008), no. 2, 153-174

[42] M. De la Iglesia, A. Durán. *Second-order differential operators having several families of orthogonal matrix polynomials as eigenfunctions.* Internat. Math. Res. Notices. IMRN 2008,Art. ID rnn 084, 24 pp.

[43] A. M. Delgado, L. Fernández, T. E. Pérez, M. A. Piñar, Y. Xu, *Orthogonal polynomials in several variables for measures with mass points*, Numer. Algorithms **55** (2010), 245-264.

[44] A. M. Delgado, L. Fernández, T. E. Pérez, M. A. Piñar, *Multivariate orthogonal polynomials and modified moment functionals*, SIGMA **12** (2016), 090, 25.

[45] M. Derevyagin, *On the relation between Darboux transformations and polynomial mappings*, J. Approx. Theory **172** (2013), 4-22.

[46] M. Derevyagin, V. Derkach, *Darboux transformations of Jacobi matrices and Padé approximation*, Linear Algebra Appl. **435** (2011), 3056-3084.

[47] M. Derevyagin, J. C. García-Ardila, F. Marcellán, *Multiple Geronimus transformations.* Linear Algebra Appl. **454** (2014), 158-183.

[48] M. Derevyagin, F. Marcellán, *A note on the Geronimus transformation and Sobolev orthogonal polynomials*, Numer. Algorithms **67** (2014), 271-287.

[49] H. Dueñas, J. García-Ardila L. Garza, A. Ramírez, *The diagonal general case of the Laguerre Sobolev type orthogonal polynomials,* Rev. Colombiana Mat. **47**, (2013), 39-66.

[50] A. J. Durán, *The Stieltjes moment problem for rapidly decreasing functions,* Proc. Amer. Math. Soc. **107** (1989), 731-741.

[51] A. J. Durán, *A generalization of Favard's theorem for polynomials satisfying a recurrence relation*, J. Approx. Theory **74** (1993), 83-109.

[52] A. J. Durán, *On orthogonal polynomials with respect to a definite positive matrix of measures*, Canad. J. Math. **47** (1995), 88-112.

[53] A. J. Durán, *Markov's theorem for orthogonal matrix polynomials,* Canad. J. Math. **48** (1996), 1180-1195.

[54] A. J. Durán, *Matrix inner product having a matrix symmetric second order differential operator,* Rocky Mountain J. Math. **27** (1997), 585-600.

[55] A. J. Durán, *Ratio asymptotic for orthogonal matrix polynomials,* J. Approx. Theory **100** (1999), 304-344.

[56] A. J. Durán, F.A. Grünbaum, *Orthogonal matrix polynomials satisfying second order differential equations*, Internat. Math. Res. Notices **10** (2004), 461-484.

[57] A. J. Durán, F. A. Grünbaum, *Orthogonal matrix polynomials, scalar-type Rodrigues' formulas and Pearson equations,* J. Approx. Theory **134** (2005), 267-280.

[58] A. J. Durán, F. A. Grünbaum, *A survey on orthogonal matrix polynomials satisfying second order differential equations,* J. Comput. Appl. Math. **178** (2005), 169-190.

[59] A. J. Durán, F. A. Grünbaum, *Structural formulas for orthogonal matrix polynomials satisfying second order differential equations,* Constr. Approx. **22** (2005), 255-271.

[60] A. Durán, M. E. H. Ismail, *Differential Coefficients of Orthogonal Matrix Polynomials*. J. Comput. Appl. Math. **190** (2006), no. 1-2, 424-436.

[61] A. J. Durán, P. López-Rodríguez, *Orthogonal matrix polynomials: Zeros and Blumenthal's theorem,* J. Approx. Theory **84** (1996), 96-118.

[62] A. J. Durán, P. López-Rodríguez, *Structural formulas for orthogonal matrix polynomials satisfying second order differential equations II*, Constr. Approx. **26** (2007), 29-47.

[63] A. Durán, V. Sánchez-Canales, *Rodrigues' formulas for orthogonal matrix polynomials satisfying second-order difference equations.* Integral Transforms Spec. Funct. **25** (2014), no. 11, 849-863.

[64] A. J. Durán, W. Van Assche, *Orthogonal matrix polynomials and higher-order recurrence relations,* Linear Algebra Appl. **219** (1995), 261-280.

[65] D. Evans L. L. Littlejohn, F. Marcellán, C. Markett, A. Ronveaux, *On recurrence relations for Sobolev polynomials*, SIAM J. Math. Anal. **26** (1995), 446-467.

[66] P. A. Fuhrmann, *Orthogonal matrix polynomials and system theory*, Rend. Sem. Mat. Univ. Politec. Torino 1987, Special Issue, 68-124.

[67] J. C. García-Ardila, L. E. Garza, F. Marcellán, *A canonical Geronimus transformation for matrix orthogonal polynomials,* Linear Multilinear Algebra. doi:10.1080/03081087.2017.1299089

[68] J. C. García-Ardila, L. E. Garza, F. Marcellán, *An Extension of the Geronimus Transformation for Orthogonal Matrix Polynomials on the Real Line* Mediterr. J. Math. **13** (2016), 5009-5032.

[69] W. Gautschi, *A survey of Gauss-Christoffel quadrature formulae,* In E. B. Christoffel (Aachen/Monschau, 1979), Birkhäuser, Basel-Boston, Mass., 1981. 72-147.

[70] W. Gautschi, *An algorithmic implementation of the generalized Christoffel theorem in Numerical Integration*, edited by G. Hämmerlin, Intern. Series Numer. Math. **57**, Birkhäuser, Basel,1982. 89-106.

[71] W. Gautschi, *Orthogonal Polynomials Computation and Approximation In Numerical Mathematics and Scientific Computation Series*, Oxford University Press, Oxford 2004.

[72] I. Gelfand, S. Gelfand, V. Retakh, R. L. Wilson, *Quasideterminants*, Adv. Math. **193** (2005), 56-141.

[73] J. S. Geronimo, *Scattering theory and matrix orthogonal polynomials on the real line,* Circuits Systems Signal Process **1** (1982), 471-495.

[74] I. M. Gel'fand, G. E. Shilov, *Generalized Functions. Volume I: Properties and Operations*, Academic Press, New York, 1964.

[75] I. M. Gel'fand, G. E. Shilov, *Generalized Functions. Volume II: Spaces of Fundamental Solutions and Generalized Functions*, Academic Press, New York, 1968.

[76] J. Geronimus, *On polynomials orthogonal with regard to a given sequence of numbers*, Comm. Inst. Sci. Math. Mec. Univ. Kharkoff [Zapiski Inst. Mat. Mech.] (4) **17** (1940), 3-18 (Russian).

[77] I. Gohberg, P. Lancaster, L. Rodman, *Matrix Polynomials*, Computer Science and Applied Mathematics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1982

[78] F. A. Grünbaum, *The Darboux process and a noncommutative bispectral problem: some explorations and challenges,* in *Geometric aspects of analysis and mechanics*, Progr. Math., *292*, Birkhäuser/Springer, New York, 2011. 161-177.

[79] F. A. Grünbaum, L. Haine, *Bispectral Darboux transformations: an extension of the Krall polynomials,* Internat. Math. Res. Notices **8** (1997), 359-392.

[80] F. A. Grünbaum, M. Yakimov, *Discrete bispectral Darboux transformations from Jacobi operators*, Pacific J. Math. **204** (2002), 395-431.

[81] F. A. Grünbaum, I. Pacharoni, J. Tirao, *Matrix valued orthogonal polynomials of the Jacobi type*, Indagat. Math. N. S. **14** (2003), 353-366.

[82] F. A. Grünbaum, I. Pacharoni, J. Tirao, *Matrix valued orthogonal polynomials of the Jacobi type: The role of group representation theory*, Ann. Inst. Fourier **55** (2005), 2051-2068.

[83] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Second Edition, Cambridge University Press, Cambridge, 2013.

[84] C. Hounga, M. N. Hounkonnou, A. Ronveaux, *New families of orthogonal polynomials*, J. Comput. Appl. Math. **193** (2006), 474-483.

[85] P. Iliev, *Krall-Laguerre commutative algebras of ordinary differential operators*, Ann. Mat. Pura Appl. (4) **192** (2013), no. 2, 203-224.

[86] A. Iserles, P. E. Koch S. P. Nørsett, J. M. Sanz Serna, *On polynomials orthogonal with respect to certain Sobolev inner products*, J. Approx. Theory **65** (1991), 151-175.

[87] A. Iserles, P. E. Koch S. P. Nørsett, J. M. Sanz Serna, *Orthogonality, and approximation in a Sobolev space. In Algorithms for approximation.* J. C. Masson and M. G. Cox, Eds. 177-124. Chapman and Hall, London, 1990.

[88] V. G. Kac, J. W. van de Leur, *The n-component KP hierarchy and representation theory*, J. Math. Phys. **44** (2003), 3245-3293.

[89] L. Karlberg, H. Wallin, *Padé type approximants and orthogonal polynomials for Markov-Stieltjes functions,* J. Comput. Appl. Math. **32** (1990), 153-157.

[90] R. Koekoek, *Generalization of the classical Laguerre polynomials and some q-analogues.* Doctoral Dissertation, Techn. Univ. of Delft, The Netherlands, 1990.

[91] R. Koekoek, H. G. Meijer, *A generalization of Laguerre Polynomials,* SIAM J. Math. Anal. **24** (1993), 768-782.

[92] E. Koelink, M. van Pruijssen, P. Román, *Matrix valued orthogonal polynomials related to* $(\mathrm{SU}(2)x\mathrm{SU}(2),\mathrm{diag})$, Internat. Math. Res. Notices (2012), **24**, 5673-5730.

[93] E. Koelink, M. van Pruijssen, P. Román, *Matrix valued orthogonal polynomials related to* $(\mathrm{SU}(2)x\mathrm{SU}(2),\mathrm{diag})$ *II*, Publ. Res. I. Math. Sci. **49** (2013), 271-312.

[94] E. Koelink, A. M. de los Ríos, P. Román, *Matrix-valued Gegenbauer polynomials*, arXiv:1403.2938.

[95] E. Koelink, P. Román, *Orthogonal vs. non-orthogonal reducibility of matrix-valued measures*, SIGMA **12** (2016) 008, 9 pages.

[96] H. L. Krall *Certain differential equations for Tchebycheff polynomials*, Duke Math. J. **4** (1938), 705-718.

[97] M. G. Krein, *The fundamental propositions of the theory of representations of Hermitian operators with deficiency index* (m,m), (Russian) Ukr. Math. J. **1** (1949), 3-66.

[98] M. G. Krein, *The ideas of P. L. Chebyshev and A. A. Markov in the theory of limiting values of integrals and their further development*, Uspekhi Mat. Nauk, 6:4(44) (1951), 3-120.

[99] K. H. Kwon, D. W. Lee, F. Marcellán, S. B. Park, *On kernel polynomials and self perturbation of orthogonal polynomials*, Ann. Mat. Pura Appl. (4) **180** (2001), 127-146.

[100] K. H. Kwon, L. L. Littlejohn, B. H. Yoo, *characterizations of orthogonal polynomials satisfying differential equations,* SIAM J. Math. Anal. **25** (1994), 976-990.

[101] I. M. Krichever, *The periodic non-Abelian Toda chain and its two-dimensional generalization* (Appendix to the paper B. A. Dubrovin, *Theta functions and non-linear equations*, Russ. Math. Surv. **36** (1981), 11-92).

[102] S. Lang, *Algebra*, Third Edition, Springer-Verlag, New York, 2005.

[103] D. C. Lewis, *Polynomial least-squares approximations*, Amer. J. Math. **69** (1947), 273-278.

[104] M. Mañas, L. Martínez Alonso, E. Medina, *Dressing methods for geometric nets: I. Conjugate nets*, J. Phys. A: Math. Gen. **33** (2000), 2871-2894.

[105] M. Mañas, L. Martínez Alonso, E. Medina, *Dressing methods for geometric nets: II. Orthogonal and Egorov nets*, J. Phys. A: Math. Gen. **33** (2000), 7181-7206.

[106] M. Mañas, L. Martínez Alonso, C. Alvárez-Fernández, *The multicomponent 2D Toda hierarchy: discrete flows and string equations*, Inverse Probl. **25** (2009), 0650077 (31pp).

[107] M. Mañas, L. Martínez Alonso, *The multicomponent 2D Toda hierarchy: dispersionless limit*, Inverse Probl. **25** (2009), 115020 (22pp).

[108] F. Marcellán, J. C. Petronilho *Orthogonal Polynomials and Coherent Pairs: The Classical Case*, Indag. Math. N.S. **6** (1995), 287-307.

[109] F. Marcellán, G. Sansigre, *On a class of matrix orthogonal polynomials on the real line*, Linear Algebra Appl. **181** (1993), 97-109.

[110] A. S. Markus *Introduction to the spectral theory of polynomials operator pencil,* Translated from the Russian by H. H. McFaden. Translation edited by Ben Silver. With an appendix by M. V. Keldysh. Translations of Mathematical Monographs, **71**. Amer. Math. Soc., Providence, RI, 1988.

[111] A. S. Markus, I. V. Mereuca, *On the complete n-tuple of roots of the operator equation corresponding to a polynomial operator bundle* Izv. Akad. Nauk. USSR Ser. Mat. **37** (1973), 1108-1131.

[112] V. B. Matveev, M. A. Salle, *Differential-difference evolution equations. II: Darboux Transformation for the Toda lattice*, Lett. Math. Phys. **3** (1979), 425-429.

[113] P. Maroni, *Sur quelques espaces de distributions qui sont des formes linéaires sur l'espace vectoriel des polynômes, in Orthogonal polynomials and their applications,* C. Brezinski et al Editors, Lecture Notes in Mathematics **1171**, Springer-Verlag, (1985), 184–194.

[114] P. Maroni, *Le calcul des formes linéaires et les polynômes orthogonaux semiclassiques, in Orthogonal polynomials and their applications*, M. Alfaro et al Editors, of Lecture Notes in Mathematics **1329**, Springer-Verlag, 1988, 279–290.

[115] P. Maroni, *Sur la suite de polynômes orthogonaux associée à la forme* $u = \delta_c + \lambda(x - c)^{-1}L$, Period. Math. Hungar. **21** (1990), 223-248.

[116] P. Maroni, I. Nicolau, *On the inverse problem of the product of a form by a polynomial: the cubic case*, Appl. Numer. Math. **45** (2003), 419-451.

[117] A. Martínez-Finkelshtein, *Asymptotic properties of Sobolev orthogonal polynomials revisited*, J. Comput. Appl. Math. **99** (1998), 491-510.

[118] H. G. Meijer *Coherent pairs and zeros of Sobolev-type orthogonal polynomials*, Indag. Math. N.S. **4** (1993), 163-176.

[119] H. G. Meijer, *Determination of all coherent pairs of functionals*, J. Approx. Theory **89** (1997), 321-343.

[120] H. G. Meijer, *A short history of orthogonal polynomials in a Sobolev space. The non discrete case*, Nieuw Arch. Wisk. **14** (1996), 93-112.

[121] A. V. Mikhailov, *The reduction problem and the inverse scattering method*, Physica D **3** (1981), 73-117.

[122] L. Miranian, *Matrix valued orthogonal polynomials on the real line: some extensions of the classical theory*, J. Phys. A. **38** (2005), 5731-5749.

[123] J. J. C. Nimmo, R. Willox, *Darboux Transformations for the Two-Dimensional Toda System*, Proc. Royal Soc. London Ser.-A 453 (1997), 2497-2525.

[124] P. J. Olver, *On Multivariate Interpolation*, Stud. in Appl. Math. **116** (2006), 201-240.

[125] A. M. Polyakov, *Quantum geometry of bosonic strings*, Phys. Lett. B **103** (1981), 207-210.

[126] A. M. Polyakov, *Quantum geometry of fermionic strings*, Phys. Lett. B **103** (1981), 211-213.

[127] A. Ronveaux, *Sobolev inner products and orthogonal polynomials of Sobolev type* Numer. Algorithms **3** (1992), 393-400.

[128] M. Rosenberg, *The square integrability of matrix valued functions with respect to a non-negative Hermitian measure*, Duke Math. J. **31** (1964), 291-298.

[129] L. Rowen, *Ring Theory,* Vol. I, Academic Press, San Diego, CA, 1988.

[130] M. A. Salle, *Darboux transformations for non-Abelian and nonlocal equations of the Toda chain type*, Theor. Math. Phys. **53** (1982), 1092-1099.

[131] L. Schwartz, *Théorie des distributions*, Hermann, Paris, 1978.

[132] B. Simon, *The Christoffel–Darboux Kernel,* in *Perspectives in partial differential Equations, harmonic Analysis and Applications*, D. Mitrea and M. Mitrea (editors), Proceedings of Symposia in Pure Mathematics **79** (2008), 295-346.

[133] A. Sinap, W. Van Assche *Polynomial interpolation and Gaussian quadrature for matrix-valued functions* Linear Algebra Appl. **207** (1994), 71-114.

[134] A. Sinap, W. Van Assche *Orthogonal matrix polynomials and applications* J. Comput. Appl. Math. **66** (1996), 27-52.

[135] J. Shohat, *On mechanical quadratures, in particular, with positive coefficients*, Trans. Amer. Math. Soc. **42** (1937), 461-496.

[136] V. Spiridonov, A. Zhedanov, *Discrete Darboux transformations, the discrete-time Toda latice, and the Askey-Wilson polynomials*, Methods Appl. Anal. **2** (1995), 369-398.

[137] V. Spiridonov, A. Zhedanov, *Discrete-time Volterra chain and classical orthogonal polynomials*, J. Phys. A: Math. Gen. **30** (1997), 8727-8737.

[138] V. Spiridonov, A. Zhedanov, *Self-similary, spectral transformations orthogonal and biorthogonal polynomials in self-similar system,* in Proc Internat. Workshop JINR, V. B. Priezzzhev, V. P. Spiridonov (Eds.), Dubna 1999, 349-361.

[139] T. J. Stieltjes, *Recherches sur les fractions continues,* Ann. Fac. Sci. Toulouse, 8 (1894), 1-122, Ann. Fac. Sci. Toulouse, 9 (1895), 1-47, English version contained in T. J. Stieltjes, Collected Papers, G. van Dijk (Ed.), Vol. II, Springer, Berlin, 1993.

[140] G. Szegő, *Orthogonal Polynomials.* 4th ed., Amer. Math. Soc. Colloq. Publ. Series, vol **23**, Amer. Math. Soc., Providence RI, (1975).

[141] J. Tirao, I. Zurrián, *Reducibility of matrix weights,* Ramanujan J. (2016). doi:10.1007/s11139-016-9834-9, 21 October 2016.

[142] K. Ueno, K. Takasaki, *Toda lattice hierarchy, in Group Representations and Systems of Differential Equations* Adv. Stud. Pure Math. **4** (1984), 1-95.

[143] V. B. Uvarov, *The connection between systems of polynomials that are orthogonal with respect to different distribution functions,* USSR Comput. Math. Phys. **9** (1969), 25-36.

[144] V. B. Uvarov, *Relation between polynomials orthogonal with different weights*, Doklady Akademii Nauk SSSR **126** (1), (1959), 33-36. In Russian.

[145] H. O. Yakhlef, F. Marcellán, M. A. Piñar, *Relative Asymptotics for Orthogonal Matrix Polynomials with Convergent Recurrence Coefficients*, J. Approx. Theory **111** (2001), 1-30.

[146] H. O. Yakhlef, F. Marcellán, M. A. Piñar, *Perturbations in the Nevai matrix class of orthogonal matrix polynomials*, Linear Algebra Appl. **336** (2001), 231-254.

[147] H. O. Yakhlef, F. Marcellán, *Relative Asymptotics for Matrix Orthogonal Polynomials for Uvarov Perturbations: The Degenerate Cases*, Mediterr. J. Math. **13** (2016), 3135-3153.

[148] G. Yoon, *Darboux transforms and orthogonal polynomials*, Bull. Korean Math. Soc. **39** (2002), 359-376.

[149] A. Zhedanov, *Rational spectral transformations and orthogonal polynomials*, J. Comput. Appl. Math. **85** (1997), 67-86.